ajer

THE ALBERTA JOURNAL OF EDUCATIONAL RESEARCH

VOLUME XXXIX NUMBER 3 SEPTEMBER 1993

PUBLISHED BY
THE UNIVERSITY OF ALBERTA • EDMONTON

The Alberta Journal of Educational Research

Published in March, June, September, and December
by the Faculty of Education, University of Alberta

A quarterly journal devoted to the dissemination, criticism, interpretation, and encouragement of all forms of systematic enquiry into education and fields related to or associated with education.

Editor: Robert H. Short

Consulting Editors

James Chapman
Massey University

Rod Clifton
University of Manitoba

Dianne Common
Lakehead University

Carol Crealock
University of Western Ontario

John Hattie
University of Western Australia

Billie Housego
University of British Columbia

Linda Larocque
University of Alberta

Stewart J.H. McCann
University College of Cape Breton

Les McLean
Ontario Institute for Studies in Education

Linda Phillips
Memorial University of Newfoundland

Sharon Robertson
University of Calgary

Faculty Advisory Committee

John G. Paterson (Chair)
Fern Snart
Joyce Edwards
Graham Fishburne

AJER gratefully acknowledges support from the Social Sciences and Humanities Research Council of Canada, from the Alberta Advisory Committee for Educational Studies, and from the Office of Research Services, University of Alberta.

The subscription rate is $32.00 per year for individuals, $40.00 per year for institutions. Add $8.00 for delivery outside Canada. Single copies are $8.00 each. Subscriptions and sales in Canada will be charged 7% GST. Please make cheques payable to *The Alberta Journal of Educational Research*. Back issues are available; rates supplied on request. Claims for undelivered copies must be received within three months of publication.

Address all communications and manuscript submissions to *The Alberta Journal of Educational Research*, Faculty of Education, Publication Services, 4-116 Education North, University of Alberta, Edmonton, Alberta, Canada, T6G 2G5.

Publications Mail Registration Number 1436

Faculty of Education
University of Alberta

The Alberta Journal of Educational Research

Volume XXXIX, Number 3, September 1993

Edwin G. Ralph **283** Sensitive, Sensible Practicum Supervision: A Contextual Application in Saskatchewan

J. Bradley Cousins
John A. Ross
Murray Prentice **297** Teachers' Evaluation of Correlational Reasoning Skills

George H. Buck **319** Faculty of Education, University of Alberta: A 30-year Gestation

Ken Pisichko **335** Manitoba's GSE: An Attempt at Resolving the Fiscal Equity Dilemma

Benjamin Levin **347** Financing Education in the Northwest Territories: Reconciling Competing Values

Diana J. Nicholson
John O. Anderson **363** A Time and Place for Observations: Talking with Primary Teachers About Classroom Assessment

M. Bryn Kulmatycki
T. Craig Montgomerie **375** A Comparison of Principals' Perceptions About Their Leadership Roles in Catholic and Non-Catholic Schools

ISSN 0002-4805

The Alberta Journal of Educational Research *Vol. XXXIX, No. 3, September 1993, 283-296*

Edwin G. Ralph
University of Saskatchewan

Sensitive, Sensible Practicum Supervision: A Contextual Application in Saskatchewan

The assistance and assessment aspects of the supervision of teaching are key responsibilities of practicum supervisors. In this article, Contextual Supervision (CS), an adaptation of Situational Leadership, is presented as a promising approach to meet both the supervisory and evaluatory goals in teacher development. The author advocates CS as a model that is not only contextually sensitive and conceptually sensible, but that has potential to assist supervisors to view globally a supervisory situation and to guide specific practice in that situation. He provides evidence to show that the CS approach (a) is consistent with goals of current reform initiatives in teacher education, (b) incorporates strengths of contemporary supervisory models, and (c) has been validated in recent supervisory practice.

Dans le domaine de la supervision de l'enseignement, la capacité de rendre service aux stagiaires et de les évaluer sont des responsabilités clés des professeur(e)s conseille(è)r(e)s associé(e)s à la pratique de l'enseignement. Dans cet article, la supervision contextuelle (Conceptual Supervison [CS]), une adaptation du leadership situationel (Situational Leadership) est présentée comme un moyen prometteur d'évaluer et de surveiller les objectifs de la formation des enseignant(e)s. L'auteur proclame la supervision contextuelle comme étant non seulement un moyen d'être sensible à la situation contextuelle et réceptif aux conditions conceptuelles, mais comme étant un moyen qui promet d'assister les professeur(e)s conseille(è)r(e)s à surveiller de façon plus globale et à diriger la pratique de l'enseignement dans cette situation particulière. Il démontre que la méthode de la supervision contextuelle (a) est conforme aux buts des initiatives de la réforme courante dans le domaine de la formation des enseignant(e)s, (b) incorpore les forces des autres modèles de supervision contemporaine, et (c) a été vérifiée lors de certains stages récents.

Introduction

A belief underlying the current reform goal of educational excellence is that assessment of teachers will improve the teaching-learning process (Porter, 1988). However, a persistent debate about assessment typically arises around such questions as: What purpose(s) will the evaluation serve? How (when and where) will it be conducted? and By whom? Generally, educators who address these questions take one of two philosophical positions delineated by Gage (1989) as the "paradigm wars." One side represents the positivistic position characterized by a technical/rational, linear approach, usually attributed to educators interested in the teaching effectiveness research (Brophy, 1992). By contrast, individuals favoring what Gage (1989) calls interpretive, qualitative, antinaturalist presuppositions are categorized at the other end of the continuum.

Edwin Ralph is in the Department of Curriculum Studies, Center for School-Based Programs in the College of Education. His specializations are supervision of teaching, professional development, and extended practicum programs.

A more realistic view, however, is one in which educators reflectively and eclectically blend both paradigms, a position that is supported by practitioners' personal experiences in the daily routine of school life, by findings in current research, and by educators' use of pragmatic common sense (Buchmann, 1990). In this article I argue that Contextual Supervision (CS) is an approach that reflects the blended position, and that has potential in helping supervisors to conceptualize the overall supervisory process and to guide their practice in the field.

CS: Description[1]

In CS, *supervision* is defined as a process of facilitating the professional self-development of educators in task-related skills. A *supervisor* is any experienced colleague (teacher or administrator) who fulfills the role of supporting or helping a partner (the *supervisee*) to meet predetermined goals established for a specific aspect of professional growth. *Assessment* refers to the evaluation of an individual's performance of a specific task or skill in order to assist that person develop professionally. An underlying assumption of CS is that assessment may effectively serve the two purposes of what has been variously described as: improvement and accountability (Glatthorn, 1990); or as supervision and evaluation (Haughey & MacElwain, 1992); or as assistance and assessment (Zimpher & Grossman, 1992). Thus the CS model agrees with what Sergiovanni and Starratt (1988, p. 351) assert: "Despite the rhetoric, everyone involved knows that evaluation is part of the process of supervision."

The Philosophy

As shown in Figure 1, the philosophical foundation of CS rests on basic human values, two of which are suggested by Frankena (1973): justice (the fair treatment of and respect for the integrity of individuals), and beneficence (the virtue devoted to the welfare of the individual[s] being served) (Sergiovanni, 1992a). When supervisors consider these values as they make supervisory decisions, then the criterion of sensitivity is addressed: supervisors respond to the unique factors related to supervisees and their particular situations.

The Principles

Rooted in this value base are CS's three guiding principles (illustrated in Figure 1) that undergird the supervisory practice. *Constructivism* conceptualizes *knowing* as a developmental process whereby all participants interactively engage in personal meaning making (Nolan & Francis, 1992). *Collaboration* is a type of community bonding that grows among the participants in supervision, in which collegiality, cooperation, and constructive confrontation emerge from the process of partners seeking to achieve the supervisory goals (Smyth, 1991). *Commitment* among participants, both to professional competence (i.e., the science, art, craft, and profession of teaching and/or supervision) and to the virtue of nurturing/caring, is the third guiding principle (Arnstine, 1990). The importance of "the caring ethic" (Greenleaf, 1977) does not imply that participants are to any degree excused from fulfilling their professional duties, but it "honors differences in teaching style and personality preferences" (Sergiovanni, 1992a, p. 117). Hersey (1985) similarly declares that effective supervisory leaders, although operating from an optimistic belief system, are obligated to confront, firmly yet positively, supervisees who may have

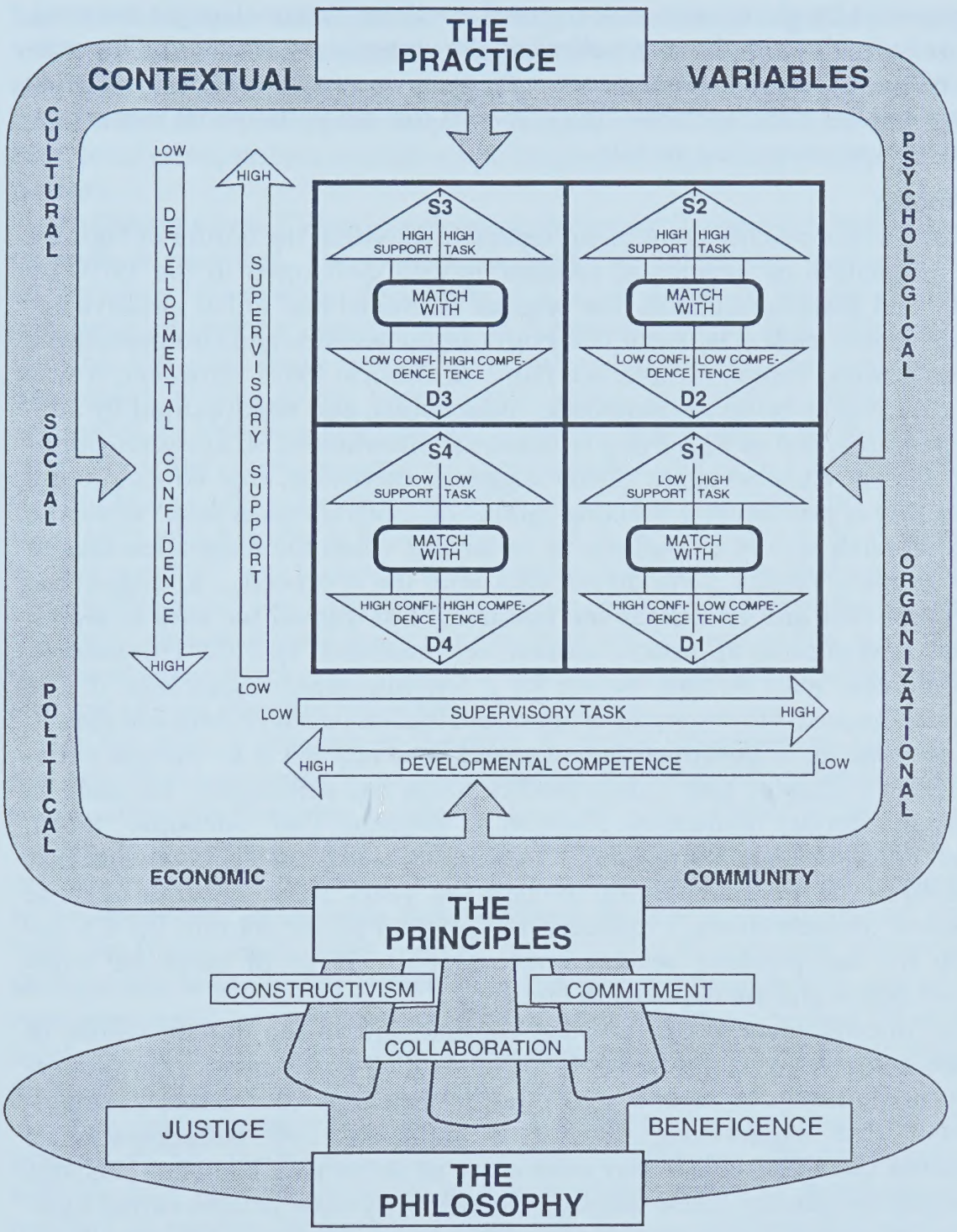

Figure 1. Contextual supervision (Blanchard, Zigarmi, & Zigarmi, 1987; Carew, Parisi-Carew, & Blanchard, 1986; Hersey & Blanchard, 1988; Niehouse, 1988a, 1988b; Ralph, 1991a, 1992, 1992-1993; Sergiovanni, 1992a, 1992b) (permission to use copyright material has been granted by the National Association of Elementary School Principals, copyright 1987. All rights reserved).

plateaued or become complacent and not shown expected growth in the task being supervised. Caring does not mean wimpiness.

The Context

The outer border in Figure 1 represents the diverse contextual variables that influence each supervisory scenario. These factors may arise from a complex web of political, economic, cultural, social, or psychological sources, often

accompanied by global, national, regional, local, institutional, organizational, and/or individual implications affecting the supervisory situation. One purpose of the CS model is not necessarily to have supervisors analyze or to attempt to alter these variables, but rather to remind participants that supervision does not occur in a vacuum.

The Practice

The heart of the practice component represented within the border of Figure 1 is an adaptation of Situational Leadership (SL) developed in the 1970s by Hersey and Blanchard (1988). The original SL model had initial intuitive appeal for school leaders in that it was both comprehensible and comprehensive (Schermerhorn, Templer, Cattaneo, Hunt, & Osborn, 1992). However, it was later criticized as being too simplistic, reductionist, and unsupported by adequate research; and as with many educational innovations, SL's original popularity waned. Kirschenbaum (1992) suggests, moreover, that educators are often guilty of prematurely rejecting promising innovations as fads: "when the initial research proves ambiguous or debatable, when the innovation fails to show dramatic results ... We do not stick with the approach ... to deepen the theory, sponsor and encourage the research," but instead we tend to search "for the next popular approach, speaker, or consultant" (p. 773). This seemed to be the case with SL and, except for a few interested researchers, it was generally ignored. However, these interested individuals who believed that SL held promise for supervisory practice, and who applied it to various situations, have suggested that certain modifications and refinements be made to the original model (Blanchard, Zigarmi, & Zigarmi, 1987; Niehouse, 1988a, 1988b). Ralph (1991b), for instance, enumerated seven such refinements that appeared in the literature during the past few years. Based on some of these suggested considerations, I reframed the original SL model into the CS approach by incorporating several improvements. Three of these are summarized below and are depicted in Figure 1.

One improvement is CS's explicit commitment to the human values of fairness and respect as guides for the actions of all participants in the supervisory relationship. A second major improvement is CS's direct acknowledgment that supervisory practice is influenced by situation-specific contextual variables, which may exist alone or in varying combinations and degrees of complexity. These unique influences may emerge from varied sources: the individual, a subgroup in the environment, the classroom, the school, the community, a region, a province, the nation, and/or an international scenario or combinations thereof. It is not the scope nor function of CS, however, to attempt to change these variables, but rather to help participants to recognize the effects of these factors and to take them into consideration in supervisory decisions.

A third key difference between SL and CS is that CS provides for a closer fit between supervisory style and supervisee development (readiness) level along the supervisory continuum. This fit consists of inverse or reciprocal relationships between both the supervisor task-supervisee competence dimension and the supervisor support-supervisee confidence dimension. Thus, in the supervision of skill-specific practice, the supervisor's task (directive) behavior is inversely proportional to the level of supervisee skill or competence (i.e., a

high level of supervisee expertise or competence in a skill requires less supervisory guidance). Simultaneously, the degree of the supervisor's support (consideration) is inversely proportional to the level of supervisee confidence (motivation or commitment) in performing a specific practice or task (i.e., a high level of supervisee confidence requires less supervisory supportive behavior).

To illustrate how CS has improved on the initial SL approach, the following example is provided. The original SL model described the continuum of a supervisee's task-specific development level in terms of learner readiness, consisting of varying combinations of supervisee ability and willingness to perform the skill. Four general points along a readiness continuum were designated, similar to the four S levels shown in Figure 1. Moreover, the original SL R1 level was described by the terms "low ability" and "low willingness" of a supervisee to engage in a particular task; but half of this designation did not fit logically with the S1 supervisory style of "low support" and "high task" with which it was to be matched. That is, the supervisory S1 dimension of "low support" was logically (and morally) inconsistent with the supervisee "low willingness" level, in that a supervisee experiencing apprehension, nervousness, or lack of confidence in a situation would need at that point a high degree of supervisory encouragement and support in addition to the other S1 element of high directive behavior.

On the basis of research on the early SL model, discrepancies such as that just described were recognized and corrected, and other adjustments were also made, although the basic quadrant structure was retained (Carew, Parisi-Carew, & Blanchard, 1986; Hersey & Blanchard, 1988). Furthermore, the original terms *readiness*, *ability*, and *willingness*, of SL were changed, respectively, to *development*, *competence*, and *confidence*; and the initial R1 and R2 supervisee readiness levels were revised, respectively, to represent supervisees' skill-specific levels of low competence and high confidence (D1), and of low competence and low confidence (D2). Consequently, these revised D1 and D2 levels synchronized more sensibly with their matching S1 and S2 supervisory styles (Blanchard, Zigarmi, & Zigarmi, 1987).

Thus, taking these refinements into consideration, I represent the actual practice of CS by the material within the border of Figure 1. The four quadrants indicate four possible positions in a supervisee's scope of development to conduct a specific instructional task. D1 (low competence or skill, but moderate to high confidence or motivation) typifies an "eager novice," such as an enthusiastic teacher intern[2] beginning the extended practicum program. D2 (low to moderate competence, and low to moderate confidence), characterizes either a "fearful neophyte" entering the field for the first time, or perhaps a "disillusioned amateur" who finds the reality of school teaching more difficult than anticipated. D3 (moderate to high competence, but fluctuating confidence), describes a "reluctant contributor" whose feelings may range from insecurity to moderate self-assurance; and D4 (high competence and high confidence), exemplifies a "peak performer." These four readiness stages are not global indicators of a person's overall development, nor are they permanent labels; and they vary among individuals and in an individual. They represent approximations in a range of ability levels (Hersey, 1985).

Supervisors seek to align their correspondingly numbered leadership style with the partner's development level: S1 matches D1, S2 with D2, and so forth. As shown in the large square in Figure 1, the four basic supervisory styles (which are also approximate points along a continuum) are:

S1 Directing. Supervisors combine high task behavior (to match their partners' low competence) and low supportive action (to correspond with supervisees' high confidence level) in order to guide, structure or "tell" the partner what is required in performing the new task.

S2 Coaching. Both high task and high supportive activities are used to bolster the supervisee's overall low to moderate development level. Specific supervisory guidance and direction are coupled with increased collaborative interaction and encouragement in helping the supervisee develop the skill.

S3 Supporting. To match the supervisee's D3 level of moderate to high competence and fluctuating confidence, S3 emphasizes the supportive more than the task element, so as to shift more authority (and responsibility) for decision making to the supervisee. The goal here is to help supervisees reduce the inconsistency in their confidence and motivation levels.

S4 Self-regulating. For supervisees functioning at a high development level, supervisors reciprocate with a leadership style that provides colleagues with considerable authority and responsibility to conduct their own self-development program. Although less supervisory direction and collaboration are required as compared with the other styles, certain contextual variables may occasionally require specific supervisor input.

CS: Rationale

The rationale supporting CS's conceptual framework is derived from three sources: (a) current reforms in teacher education, (b) three contemporary supervisory models, and (c) recent validation of CS in actual supervisory practice.

CS and Current Reform Goals

The philosophical base and the 3C principles of CS represented in Figure 1 support recent educational reform goals referred to earlier in this article. The first C, Constructivism, which is one of the key elements in the reform movement (Leinhardt, 1992; Willis, 1992), is central to CS. Cognitive developmental psychology has informed us that individuals, including those in supervisory and supervisee roles, learn by actively engaging in personal meaning making, as they seek solutions to real problems encountered in their daily life routines. CS has been shown to be effective in assisting supervisory participants in this quest (Ralph, 1992). The second C principle in CS, Collaboration, occurs through interaction among team members as they assume various roles in the supervisory process, sometimes on a reciprocal, exchange, or rotational basis, depending on their unique experiences/ skills/confidence levels in specific instructionally related tasks. Collaboration in CS would include cooperation and collegiality, as well as confrontation, between or among team members. However, the key is that these interactions would all be based on the underlying values of justice and beneficence that guide participants as they pursue the main goal of mutually assisting partners to meet agreed-on supervisory objectives.

Commitment, the third C principle supporting CS practice, refers to individual and team responsibility for achieving an appropriate balance between the dual goals of improving professional competence (i.e., the task component) and of demonstrating personal concern, support and encouragement (i.e., the human relationship element). The key here is that each supervisory situation is unique: leadership style must be adjusted to each context.

CS and Other Supervisory Approaches

The rationale supporting CS incorporates the strengths and avoids the limitations of three prominent supervisory approaches. The practice component of CS, shown within the border of Figure 1 as an adaptation of *Situational Leadership* (Hersey & Blanchard, 1988; Ralph, 1991b, 1992-1993), shares SL's advantages: it emphasizes adaptive leadership rather than "one best method"; it is appealing because it is relatively easy to understand (Caskey, 1988; Hersey, 1985); and it helps bridge the theory-practice gap by clarifying the conceptualization of individual supervision and by offering practical supervisory guidance (Niehouse, 1988b).

SL (and thus CS) has been criticized for being reductionist, because it is said to concentrate solely on the synchronization of supervisor style and supervisee development at the expense of a host of other situational factors influencing each context. Although CS does in fact emphasize this style-level match, it does so because the matching process is considered to be the heart of supervision. Furthermore, CS not only acknowledges the effect of contextual influences, but it helps alert the supervisor to consider them in supervisory practice. Yet the purpose of CS is not to overemphasize these variables to the extent that supervisors would become preoccupied with modifying individuals' personalities, cultural characteristics, school-community relations, or institutional policies. Indeed, individuals in supervisory roles are realistically only able to concentrate on, and have responsibility for, changing one important element over which they do have any personal control, and that is to adjust their own leadership style. They are not expected, physically or ethically, nor do they have the resources and authority to attempt, to manipulate other people, practices, or policies. In addition, many of these contextual factors may not even be addressable by participants during the daily routines of supervisory practice in schools (Blanchard, Zigarmi, & Zigarmi, 1987). Thus the sensible view of CS as a guide for helping supervisors to change only what they are able to change is more fitting with reality than the view that conceptualizes CS as merely reducing supervision to simplistic or prescriptive behaviors. The former view is not only more intuitively appealing, but it has been shown to be effective in guiding, but not restricting, supervisory decisions (Cawelti, 1976; Gates, Blanchard, & Hersey, 1976; Ralph, 1992; Zigarmi, 1981).

A second supervisory approach having some relation to CS is *Developmental Supervision* (DVS, Glickman, 1990). Both CS and DVS have roots in the SL model, both emphasize individuals' self-development, and both follow an interactive process in determining the fit of development level with supervisory style. They differ in some areas, however. One is that DVS is interventionist, in that supervisors have the responsibility of diagnosing and prescribing what they decide supervisees need. Although CS agrees with DVS in assuming that prescriptive structure may be required on occasion, CS is

more comprehensive than DVS in that it also provides for supervisee-regulated decisions at the D4 level. A second difference is that DVS concentrates on supervisees' levels of cognitive or abstract thinking development, whereas CS accepts a broader range of indicators of development (e.g., intuition and connoisseurship). A third difference is the DVS has only one middle development category, whereas CS has two. CS thereby encourages more refined determination of supervisee skill level by supervisors in selecting the more effective combination of supervisory task and support behaviors in quadrants 2 and 3.

A third current model, *Differentiated Supervision* (DFS, Glatthorn, 1990), also shares some commonalities with CS. For instance, both plans are developmental and are less prescriptive than Glickman's (1990) DVS. A main difference between CS and DFS, as was the case with DVS, is CS's provision for four rather than three supervisory styles, which provides for more accurate determination of supervisee skill level and synchronization of supervisory action within the two middle categories. Moreover, rather than providing complete supervisee choice of the type of supervision they want, CS retains the high task (directive) mode in S1 and S2 because certain supervisees may not yet have developed the expertise to select the type of supervision suited for a particular task and thus may require specific supervisory guidance.

Critics may argue that it is unrealistic to expect supervisors to master the four CS leadership styles, and that it is simply too difficult for them to change their preferred style or personality characteristics. In response, I would suggest that educators, by virtue of the "people profession" they are in, have learned to be adaptive and to be actors, often having to adjust their behavior and feelings to accommodate unanticipated, unnatural, and/or unpleasant circumstances. In this vein, Fullan (1991) reports that educators will work at incorporating an innovation provided they see it as having practical value in improving their own practice. I believe this assertion to be valid in the case of CS, and I provide evidence to that effect in the next section of this article. This evidence represents the third source of the rationale supporting the CS model.

Implementation

The application of the CS approach in actual supervisory practice first would require that the goal of the supervisory situation be clarified, this usually being the improvement of one or more of the supervisee's professional skills. Second, the supervisee's task-specific development level for this skill would be ascertained through the pair's face-to-face dialogue, questioning and/or observation by the supervisor. During this interaction, the supervisor's professional obligation would be to emphasize openness, trust, and mutual respect, and the pair would no doubt also address some of the contextual factors impinging on their situation. Supervisors would then determine the leadership style most appropriately fitting their partners' readiness levels and would subsequently adjust both their directive and relationship actions (Rust, 1988).

Assumptions

In applying the CS approach in my own supervisory practice, I took into consideration the following assumptions (Blanchard, Zigarmi, & Zigarmi, 1987; Niehouse, 1988a, 1988b): (a) the spirit behind the CS approach is to offer suggestions and guidance for supervisory decisions, not to prescribe absolute rules or directions; (b) each of the four development levels and supervisory

styles described in the model represents the probable range of behavior connected with specific combinations of criteria: they are not rigid categorizations or unchangeable characteristics; and (c) the alignment of style and readiness occurs within a leadership zone—a fluid range of supervisory action that is sensitive to nuances of personalities among participants.

Actual Application

From 1990 to 1992 in my position as Supervisor of Extended Practicum with the University of Saskatchewan, I applied the CS model in my supervisory relationships with 49 pairs of teacher interns[2] and classroom cooperating teachers (CCTs) during the 16-week internship in several schools throughout the province. I applied the CS framework as described in the previous section both during formal and informal supervisory conferences with participants, and during personal periods of reflection-on-action (Schön, 1987). For approximately 205 supervisory episodes, I wrote field notes, selectively recording key observations, descriptions, incidents, and evaluatory comments related to each situation. However, because of space limitations, I report here only a summary of my CS relationships as they occurred with the 49 pairs at a specific point in the practicum. The summary is like a composite photo of 49 still frames taken during the 14th week of the practicum, which shows an average of the interns' level of presenting/structuring skills in formally observed lessons[3] (University of Saskatchewan, 1992). Extending this photograph metaphor, one could envisage ending up with a blurred image. This possibility illustrates the methodological danger of attempting to force the blending of qualitative and quantitative paradigms.

During my fourth of five supervisory visits with each CCT-intern pair during the extended practicum period, I concentrated on the area of effective presentation skills,[3] which was one of several major areas of teaching emphasized throughout our practicum program (e.g., professional attributes, lesson planning, unit planning, questioning, classroom management). Each of these major categories had several performance indicators that interns were expected to internalize and demonstrate during their practicum teaching.

During the fourth visit, I assessed each intern's development level in this presenting category and also determined the general supervisory style characterizing the CCT's interaction with the intern at that time. Having met the pairs in their schools on three earlier occasions, as well as at three large-group practicum seminars, I was able to compare the teacher interns' development levels and CCTs' leadership styles with their respective levels and styles observed previously. I believe that my assessment skills in this process were bolstered by my seven years of past supervisory experience with teacher interns and seasoned practitioners. This prior background provided me with both confidence and skill in applying the CS model as a framework to guide my own supervisory observations, questions, interviews, conversations, and recording of data with individuals and pairs. On the basis of the data collected through these procedures I was able to classify[4] interns' presentation skill levels and CCTs' leadership styles according to the four pairs of supervisory categories in the CS model during the 14th week of the internship. I then calculated the percentages of participants who were in each category and

summarized these findings as shown in Figure 2. However, before interpreting these results, I delineate three limitations related to this phase of the study.

First, the CS approach is designed for supervisory practice at the micro level of individual development rather than the macro organizational level. Thus the quantifying or averaging of group performance is not part of the CS framework, but only reflects my attempt to determine the trends, if any, that may have emerged from an analysis of all of the supervisory scenarios as a whole. A second caveat is that none of the participants had knowledge of the CS model or of my research intentions; thus two possible strands of interpretation are possible for this section of the study. One strand relates to the findings viewed from my perspective as a college supervisor seeking to achieve the supervisory and evaluatory goals of the supervisor's position. The other, based on an analysis of the group supervisory practice of 49 CCTs who were unfamiliar with CS, has implications for the possible implementation of CS training for future practicum supervisors and/or CCTs.

A third condition affecting the interpretation of the data in Figure 2 is that these findings reflect only one observer's judgments. Yet, in keeping with current trends toward school-based, collaborative action research, such observations may contribute to the knowledge in the field.

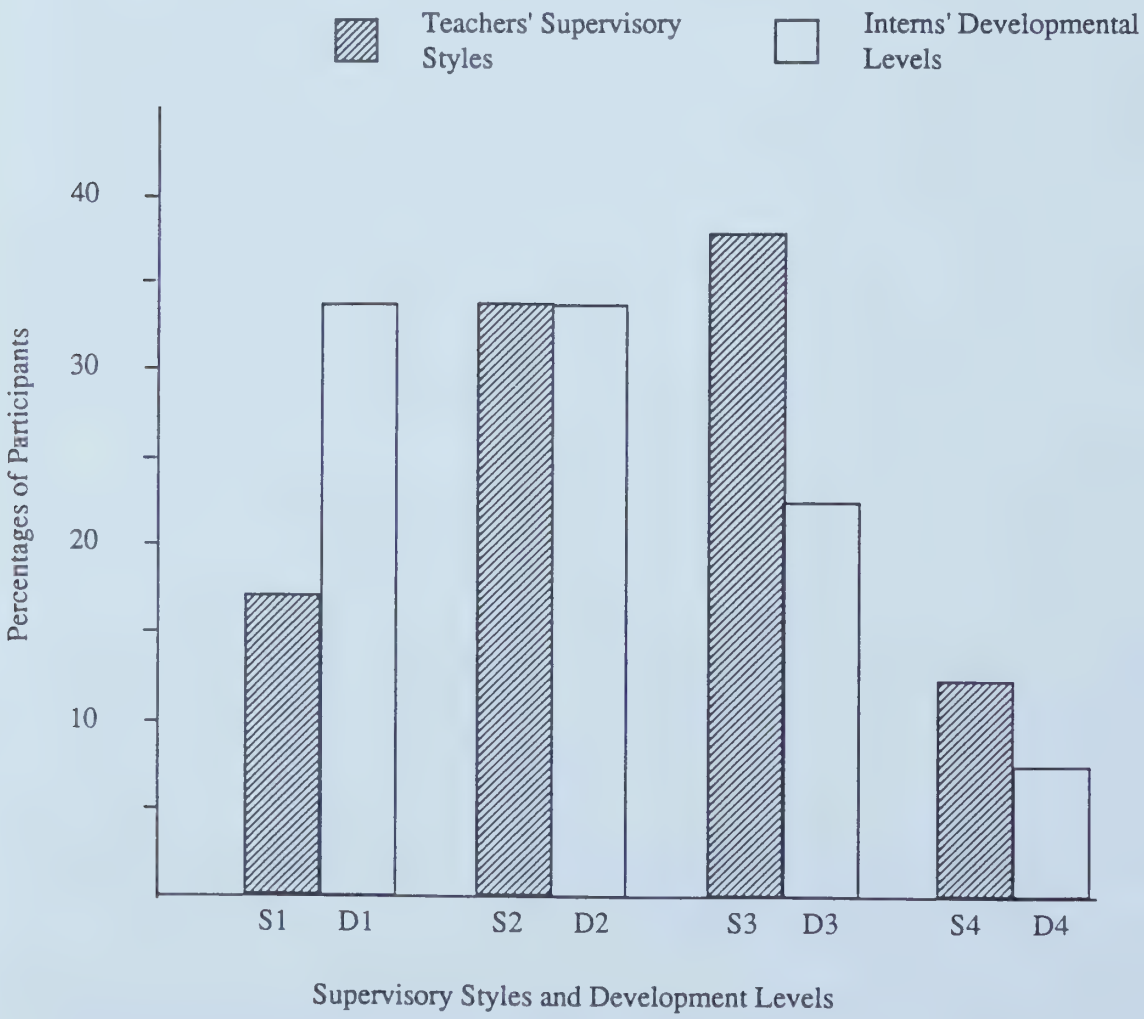

Figure 2. Degree of match between teachers' supervisory styles and interns' development levels for lesson presenting skills determined during 14th week of 16-week extended practicum. N=49 supervisory pairs. None had CS training.

Interpretation and Discussion of Findings

Taking into consideration the above limitations, I discuss the findings from two perspectives: (a) an analysis of practicum supervision among 49 pairs untrained in the CS approach, and (b) my experience of applying CS to my own supervisory practice. With respect to the first perspective, I summarize the qualitative data in Figure 2 and show that at a specific point during the 14th week of the practicum, there were certain percentages of participants who were at various stages in the supervisory process. The CCTs are represented by the shaded bars and the interns by the blank bars.

The horizontal axis represents the combinations of matches between supervisory style and supervisee development level in instructional presenting skills. Ideally, S1 would match with D1, S2 with D2, and so on. The vertical axis represents percentages of participants who were at specific stages during the 14th week of the practicum. The shaded bars indicate approximate percentages of supervisors who were exhibiting a particular leadership style in their supervisory relationship with their teacher interns as the pairs worked on presenting/structuring skills. The similarly numbered blank bars represent the percentages of interns who were at those skill development levels during the 14th week.

An examination of Figure 2 reveals that there was little synchronization of supervisory style with supervisee development level, a finding to be expected because supervisors were not knowledgeable about CS. A second general observation, confirming similar patterns indicated in past research (Glatthorn, 1990; Zeichner, 1992), is that although the 49 teacher interns were at different development levels during the 14th week, the majority were at the two lower levels (34% at D1, 34% at D2), with progressively fewer at D3 (23%) and D4 (8%). Third, one sees from Figure 2 that the second level was the only point at which supervisors as a group matched their S2 (coaching) style with an equal number (34%) of interns at the D2 level (moderate competence and confidence). This is consistent with other research that indicates that teacher supervisors generally prefer a style that is high in both human and task elements (Glatthorn, 1990). Moreover, this tendency among supervisors to avoid a directive style may be further illustrated by comparing the number of supervisors using the S1 style (17%) with the number of teacher interns at the D1 level (34%). Similarly at both the third and fourth stages, more supervisors (37% and 13% respectively) appear to favor the supportive, human dimension of supervision (S3 and S4) than were required by the lower percentages of supervisees at the D3 and D4 levels (23% and 8% respectively).

Thus an implication for practicum supervision begins to emerge from these findings, according to CS theory: that the general mismatch between supervisory style and supervisee development level could be corrected if supervisors would implement more directive supervisory action when interns' readiness levels so require.

With respect to the second strand of data interpretation based on the perspective of one college supervisor using the CS model in 205 supervisory situations, I offer the following insights. First, I found the CS model to be a supervisory tool capable of being used in a reflective manner to help accomplish the ultimate goal of improving the teaching-learning process for all participants: pupils, interns, CCTs, and college personnel. I applied CS in the

manner described in the "Implementation" section above as a guide to help (a) orient myself to the contextual mosaic unique to each supervisory situation; (b) conceptualize the supervisory process for each scenario; (c) determine the supervisee readiness level for each instructional task; and (d) decide what combination of task and support behaviors I needed to use in order to meet supervisees' levels of competence and confidence.

I did not follow the model prescriptively but used it reflectively and selectively before, and/or during, and/or after formal supervisory conferences and informal conversations with participants. Because my personal philosophy is consistent with the values of justice and beneficence as espoused by CS, I did not feel guilty about being flexible with and sensitive to the pairs in determining the mix of supervisory direction and support appropriate for each task. Moreover, I was comfortable using CS's sensible guidelines, which allowed me to take into consideration the particular personal, organizational, and cultural factors characterizing each pair's context. For reports of specific CS scenarios, see Ralph (1991a, 1992, 1992-1993).

Furthermore, I believe that through my application of CS both the supervisory and evaluatory roles of my position were fulfilled effectively. Although I endeavored to establish and maintain principles of clear communication, fairness, and trust in each triad relationship, I found that the majority of the pairs initially expressed some degree of concern about the intern evaluation process. However, by the end of the practicum, they showed satisfaction with the assessment procedures, as revealed by participants' written evaluatory comments on the college supervisor's work (an evaluation required by the University). Several of these comments indicated that respondents viewed the intern evaluation procedure as an integral component of the supervisory function, and that when participants regarded positively the college supervisor's overall performance, then a similar view was held for the performance assessment process of the interns' teaching. It appeared that acceptance of the supervisor was directly related to acceptance of the assessment process.

Implications and Conclusions

Based on my personal application of the model from 1990 to 1992, I found CS to be useful in clarifying my conceptualization of the supervisory contexts; in guiding specific supervisory practice in these situations; in helping to establish trust during the supervisory process; and in helping to achieve the fundamental goal of supervision and evaluation: to help educators develop professionally so that the teaching-learning process in schools will be enhanced.

With respect to the findings related to the supervisory episodes among the 49 pairs, an implication for our practicum program is that providing a supervisors' inservice program on CS for CCTs would in all probability help reduce the mismatch often arising between supervisory styles and readiness levels. If all our practicum supervisors were knowledgeable of the conceptual model as a means of promoting sensible supervisory guidance, then they would be equipped to determine more accurately their partners' readiness levels and to tailor their supervision appropriately by adjusting leadership behavior. Implementing CS is superior to clinging inflexibly to a single supervisory style or to operating unsystematically or erratically via a trial-and-error approach. Familiarity with the model would assist supervisors who may

misinterpret others' lack of skill or confidence as resistance or stubbornness to reframe these normal responses into legitimate problems to address and solve during the supervisory process (Fullan & Miles, 1992, p. 748).

Believing that CS has merit, our College has, at the time of this writing, set up an orientation workshop on the model for faculty members supervising in the extended practicum. The plan is for these individuals to learn to apply CS in their supervisory practice in a manner similar to my experience described in this article, and then on the basis of the group's experience, to determine whether or not to broaden this CS inservice to include all future practicum supervisors, both teachers and faculty members.

Critics suggest that contingency leadership theories, which are similar to CS, will lead to naive prescription applied to practice (Sergiovanni & Starratt, 1988, p. 213); or that any positive results arising from CS's application could just as easily have been achieved without the use of the model. However, in the light of the findings presented in this article, I believe that such criticisms are unwarranted. Our recent experience in Saskatchewan provides compelling evidence that CS has potential to inform supervisory practice and to add precision to making sense of participants' actions, including one's own.

Notes

1. The CS model has been described in other articles by Ralph (1991a, 1992, 1992-1993).
2. In this article the terms *intern, teacher intern,* and *student teacher* are synonymous. Similarly, the terms *internship, extended practicum,* and *practicum* all represent the 16-week extended practicum program for all BEd students at the University of Saskatchewan.
3. *The Internship Manual* (University of Saskatchewan, 1992, p. 66) refers to "presenting skills" as "the effective delivery of the planned lesson ... dependent on clarity and enthusiasm." The manual describes 12 instructional skills derived from current teacher effectiveness research that are indicative of clear presentation (e.g., "Clearly states objectives," "Gives clear directions and explanations," "Makes clear transitions," and "Provides summaries").
4. Further research could investigate the reliability and validity question of this assessment process by having two or more observers trained in CS to evaluate (in person or by videotape) the same supervisory scenario(s) and then to compare their evaluations.

References

Arnstine, B. (1990). Rational and caring teachers: Reconstructing teacher preparation. *Teachers College Record, 92*, 230-247.

Blanchard, K., Zigarmi, D., & Zigarmi, P. (1987). Situational leadership: "Different strokes for different folks." *Principal, 66*(4), 12-16.

Brophy, J. (1992) Probing the subtleties of subject-matter teaching. *Educational Leadership, 49*(7), 4-8.

Buchmann, M. (1990). Review of *Teachers' Professional Learning*. *Journal of Teacher Education, 41*(4), 54-59.

Carew, D., Parisi-Carew, E., & Blanchard, K. (1986). Group development and situational leadership: A model for managing groups. *Training and Development Journal, 40*(6), 46-50.

Caskey, F. (1988). *Leadership style and team process: A comparison of the managerial grid and situational leadership.* (Training and Development Research Center Project No. 25). St. Paul, MN: Minnesota University, Department of Vocational and Technical Education. (ERIC Document Reproduction Service No. ED296 162)

Cawelti, G. (1976). *Selecting appropriate leadership styles for instructional improvement* (Videotape). Alexandria, VA: Association for Supervision and Curriculum Development.

Frankena, W. (1973). *Ethics*. Englewood Cliffs, NJ: Prentice-Hall.

Fullan, M. (1991). *The new meaning of educational change* (2nd ed.). New York: Teachers College Press.

Fullan, M., & Miles, M. (1992). Getting reform right: What works and what doesn't. *Phi Delta Kappan, 73*, 744-752.

Gage, N. (1989). The paradigm wars and their aftermath. *Educational Research, 18*(7), 4-10.

Gates, P., Blanchard, K., & Hersey, P. (1976). Diagnosing educational leadership problems: A situational approach. *Educational Leadership, 33*(5), 348-354.
Glatthorn, A. (1990). *Supervisory leadership: Introduction to instructional supervision.* Glenview, IL: Scott, Foresman/Little, Brown.
Glickman, C. (1990). *Supervision of instruction: A developmental approach* (2nd ed.). Boston, MA: Allyn and Bacon.
Greenleaf, R. (1977). *Servant leadership.* New York: Paulist Press.
Haughey, M., & MacElwain, L. (1992). Principals as instructional supervisors. *Alberta Journal of Educational Research, 38,* 105-119.
Hersey, P. (1985). *The situational leader.* New York: Warner.
Hersey, P., & Blanchard, K. (1988). *Management of organizational behavior: Utilizing human resources* (5th ed.). Englewood Cliffs, NJ: Prentice-Hall.
Kirschenbaum, H. (1992). A comprehensive model for values education and moral education. *Phi Delta Kappan, 73,* 771-776.
Leinhardt, G. (1992). What research on learning tells us about teaching. *Educational Leadership, 49*(7), 20-25.
Niehouse, O. (1988a). Leadership concepts for the principal: A practical approach (Part 1). *NASSP Bulletin, 72*(504), 50-60.
Niehouse, O. (1988b). Leadership concepts for the principal: A practical approach (Part 2). *NASSP Bulletin, 72*(505), 62-71.
Nolan, J., & Francis, P. (1992). Changing perspectives in curriculum and instruction. In C. Glickman (Ed.), *Supervision in transition: The 1992 ASCD yearbook* (pp. 44-60). Alexandria, VA: Association for Supervision and Curriculum Development.
Porter, A. (1988). Understanding teaching: A model of assessment. *Journal of Teacher Education, 39*(4), 2-7.
Ralph, E. (1991a). Aboriginal student-teachers in the extended practicum: Contextual supervision. *Teaching Education, 4*(1), 113-121.
Ralph, E. (1991b). *Supervising the extended practicum: A situational approach.* Unpublished manuscript, University of Saskatchewan, College of Education.
Ralph, E. (1992, September). *Continuous teacher development: A case for contextual supervision.* Paper presented at the annual conference of the British Educational Management and Administration Society, Bristol, UK. (The text of this presentation subsequently appeared in *Brock Education, 2*(3), 1992, 13-17.
Ralph, E. (1992-1993). Improving practicum supervision: A Canadian experience. *Action in Teacher Education, 14*(4), 30-38.
Rust, F. (1988). How supervisors think about teaching. *Journal of Teacher Education, 39*(2), 56-64.
Schermerhorn, J. Jr., Templer, A., Cattaneo, J., Hunt, J., & Osborn, R. (1992). *Managing organizational behavior.* Toronto: Wiley.
Schön, D. (1987). *Educating the reflective practitioner: Toward a new design for teaching and learning in the professions.* San Francisco, CA: Jossey-Bass.
Sergiovanni, T. (1992a). *Moral leadership: Getting to the heart of school improvement.* San Francisco, CA: Jossey-Bass.
Sergiovanni, T. (1992b). Why we should seek substitutes for leadership. *Educational Leadership, 49*(5), 41-45.
Sergiovanni, T., & Starratt, R. (1988). *Supervision: Human perspectives* (4th ed.). New York: McGraw-Hill.
Smyth, J. (1991). Instructional supervision and the redefinition of who does it in schools. *Journal of Curriculum and Supervision, 7*(1), 90-99.
University of Saskatchewan. (1992). *The internship manual.* Saskatoon: College of Education, Center for School-Based Programs.
Willis, S. (1992, June). Teaching thinking. *Association for Supervision and Curriculum Development Curriculum Update,* pp. 1-8.
Zeichner, K. (1992, February). *Beyond inquiry-oriented teacher education: Rethinking the practicum in North American teacher education.* Paper presented at the annual meeting of the Western Canadian Association for Student Teaching, Edmonton.
Zigarmi, D. (1981). Leadership and school climate: A data based approach to administrative training. *Journal of Staff Development, 2*(1), 93-115.
Zimpher, N., & Grossman, J. (1992). Collegial support by teacher mentors and peer consultants. In C. Glickman (Ed.), *Supervision in transition: The 1992 ASCD yearbook* (pp. 141-154). Alexandria, VA: Association for Supervision and Curriculum Development.

The Alberta Journal of Educational Research *Vol. XXXIX, No. 3, September 1993, 297-317*

J. Bradley Cousins
University of Ottawa

John A. Ross
Ontario Institute for Studies in Education

and

Murray Prentice
Durham Board of Education

Teachers' Evaluation of Correlational Reasoning Skills

Two experiments sought to increase our understanding about teachers' proficiency in assessing higher order thinking skills. In Experiment 1 eight grade 7 teachers participated in training and subsequently assessed samples of student performance from grades 5, 7, 9, 11, and OAC (Ontario Academic Credit) on two forms of a correlational reasoning test. Teachers' ratings of the 100 samples of student performance did not correspond highly with an expert's assessment although the novice assessors did show sensitivity to variation in student performance due to student age and test difficulty. In Experiment 2, 31 grade 7 and 8 teachers participated in an inservice program on correlational reasoning that required them to implement an instructional unit in their own classrooms. They were subsequently trained to assess student performance and marked a sample of pre- and posttest data from their own students. The prediction that more intense implementation would enhance teachers' understanding of correlational reasoning concepts and concomitantly, their assessment skills was not supported. Teachers' assessment proficiency was higher than in Experiment 1, but did not correspond highly with the standard set by the expert marker. Sensitivity to student growth was evident, however. Results are discussed in terms of teachers' purposes for doing assessment and implications for research.

Cet article présente les résultats de deux expériences qui cherchaient à nous aider à mieux comprendre la capacité qu'ont les enseignant(e)s à évaluer les habiletés cognitives supérieures des élèves. Dans la première expérience, Expérience 1, huit enseignant(e)s de la septième année ont été formés à pouvoir déterminer le niveau de performance des élèves des niveaux cinquième, septième, neuvième et onzième années ainsi que de l'OAC (Ontario Academic Credit) sur deux formulaires d'un test qui servait à mesurer le raisonnement correlationnel. L'évaluation des enseignant(e)s de 100 exemplaires de travaux d'élèves ne correspondait pas tellement avec l'évaluation d'un expert même si les assesseurs novices ont démontré une sensibilité aux variations individuelles des élèves en raison de leurs âges et de la difficulté des deux tests. Dans la deuxième expérience, Expérience 2, 31 enseignant(e)s de la septième et de la huitième année ont participé à un atelier de formation sur le raisonnement correlationnel qui les requérait l'introduction d'une unité d'étude dans leurs salles de

---

Brad Cousins is an associate professor of education. His interests are in program evaluation and student assessment. He was located at the OISE Trent Valley Centre when the current work was carried out

John Ross is a professor of curriculum and head of the Trent Valley Centre. He specializes in curriculum development, cooperative group learning, and program evaluation.

Murray Prentice is a director of research. His interests are in teaching higher order thinking skills and program evaluation.

classes respectives. Subséquemment, on les a formé à évaluer la performance de leurs élèves en évaluant des données de pré et post-tests tirées de leurs propres élèves. La prédiction qu'une mise en oeuvre plus judicieusement exécutée améliorerait la compréhension des enseignants(e)s vis-à-vis le raisonnement correlationnel et conséquemment leurs habiletés d'évaluer étaient supérieures à celles de l'Expérience 1. Cependant, cette meilleure capacité d'évaluer ne correspondait pas avec les normes établies par l'évaluateur expert. En plus, la sensibilité au cheminement des élèves a été mis en évidence. Les résultats de ces deux expériences sont discutés en fonction des objectifs de l'évaluation pour les enseignant(e)s et en fonction des implications pour de plus amples recherches à ce sujet.

Advocates of educational reform are united in their belief that change in classroom practice is key. But such advocates also agree that fundamental change in instruction depends quite seriously on reform in the assessment of student performance (Linn, Baker, & Dunbar, 1991; Shepard, 1991). Research on change in assessment practice, however, has been limited. Only a few studies have examined the effects of intervention programs designed to enhance teachers' assessment proficiency, although the need for such programs has been clearly underscored (Gullickson, 1984; Rogers, 1991; Schafer & Lissitz, 1987; Stiggins, 1988).

Abundant research has accumulated in recent years concerning teachers' student evaluation practices. In contemporary reviews of empirical research (Crooks, 1988; Earl & Cousins, 1993) several practice-oriented issues, problems and breaches in our knowledge concerning student evaluation have been identified. The majority of the questions on teacher-made tests, for example, assess low level cognitive knowledge of terms, facts, rules, and principles, but very few assess behaviors that would be classified as requiring higher order skills such as application, synthesis, or evaluation (Bateson, 1990; Crooks, 1988; Fleming & Chambers, 1983; Rogers, 1991; Stiggins, Frisbie, & Griswold, 1989; Wilson, 1990). One possible explanation for this observation is that low level cognitive outcomes are easier and more economical to measure and as such fit better with teachers' fragmented daily schedules. Our knowledge of teachers' ability to measure validly and reliably higher order thinking skills is limited and ought to be expanded.

A second concern is that although the plea for assessment techniques that are referenced to specific curriculum has been resounding, evidence suggests that although such practices may be relatively prosaic in elementary grades, they become much more heavily referenced toward generalized achievement in the senior grades (McLean, 1985). However, when standards are outcome-referenced and assessment is aligned with curriculum objectives and instruction student achievement improves significantly (Bloom, 1984). Also, motivational consequences of using individual-referenced standards are salubrious (McIver, 1989; Natriello, 1987; Slavin, 1980), but constraints imposed on teachers by school or provincial policy, parental requirements, and lack of time have been severely limiting (Nagy & Moorehead, 1990; Stiggins et al., 1989; Wilson, 1989; Wilson, Rees & Connock, 1989). In addition to teacher-constructed assessment tools, many studies have shown strong teacher preferences for the use of projects, experiment writeups, and written work as a basis for evaluation (Anderson, 1989; Bateson, 1990; Dorr-Bremme, 1983; Stiggins, 1988; Wilson, 1990). Indeed, several authors have provided compelling arguments for moving toward direct or "authentic" modes of assessment (McLean,

1990; Wiggins, 1989). Researchers caution, however, that we need to subject assessment alternatives to the same validity investigations as traditional tests (Linn et al., 1991; Shepard, 1991) and that moving too quickly in this direction would not be prudent (Shavelson, Baxter, & Pine, 1992). Teachers, especially in lower grades, do engage in performance-based approaches to assessment (Dorr-Bremme, 1983; Stiggins & Bridgeford, 1985) but structured performance assessment methods that are more appropriate for assessing higher levels of skill are notably absent in the senior elementary grades and beyond. We need to investigate the viability of using such methods at the senior level.

A third concern about teachers' assessment practice has to do with time management issues. Teachers spend considerable time and energy evaluating student performance. Stiggins (1988) found that teachers typically spend about one third or more of their professional time in assessment-related activities. Wilson (1989) found variability from class to class, but suggested that high school students carrying a full program could expect to engage in evaluative activities every second day. But even with such time earmarked for assessment activities, the bulk of teachers' effort is invested in marking and grading rather than developing new assessment instruments or evaluation procedures (Stiggins & Bridgeford, 1985). In addition, modest changes or additions to instructional program can place severe limits on teachers' ability to credibly evaluate students. Sizer (1984), for example, observed that a high school English teacher with responsibility for 120 students per day is faced with 10 hours per week of added marking for every additional student assignment requiring only five minutes to evaluate. Teachers do not have an abundance of time to develop new evaluation instruments nor even their own proficiency in evaluating students.

Fourth, teachers tend to grade student work in isolation from their staff peers (Wilson, 1989) and to rely heavily on self-developed tests and evaluation instruments. By the time students reach the senior elementary grades, marks are determined substantially by teacher-made tests, and in secondary schools test scores are equated with achievement (McLean, 1985). However, there are few data to show that teacher-made tests are either valid or reliable by conventional standards. At the opposite end of the continuum, in the United States more than Canada, reliance on externally produced standardized, norm-referenced tests of student achievement is quite high, although the utility of such tests for classroom purposes is not clear (Paris, Lawton, Turner, & Roth, 1991; Salmon-Cox, 1981; Smith, 1991; Sproull & Zubrow, 1981). Paris et al. (1991), for example, identified several negative consequences for students of large-scale testing programs and Smith (1991) listed several consequences for teachers. In Canada, Bateson (1990) found that British Columbia teachers would prefer school-wide tests to provincial or even district-wide assessment of student achievement. However, preferences were still for tests referenced to generalized achievement. Bongar (1982) found that grade 6 teachers in Alberta rejected Canadian Test of Basic Skills (CTBS) scores that did not match their expectations, regardless of whether student performance was above or below what they expected. Airasian (1979) found that after repeated exposure to external test results, teachers were more likely to use them for making grouping decisions, but they still saw these test scores as less accurate than other data

such as their own observations. Herman and Dorr-Bremme (1983) reviewed evidence to suggest that teachers pool the results of externally produced tests with data produced by their in-class observations and teacher-made tests. The external tests may have less influence than some would like because their impact is diluted with other instruments by teachers who are indifferent to psychometric adequacy (Dorr-Bremme, 1983). Stiggins (1987) developed a module for teachers to assist in developing their own structured performance assessment instruments. Given the time constraints on teachers, under some circumstances it may be more feasible to develop externally valid and reliable criterion-related tests of student performance. We need to know the extent to which teachers are able to successfully implement such externally produced assessment systems.

A fifth concern is that fundamental measurement theory is missing in the classroom (McLean, 1985; Schafer & Lissitz, 1987; Stiggins & Bridgeford, 1985; Wilson, 1989; 1990). Teachers tend not to focus on the assessment of students' characteristics when they plan, nor on goals and objectives. Rather, the instructional activity is most often the planning unit. Stiggins and colleagues suggested that teachers' evaluation practices lack rigor and possibly appropriateness and that the impact of consequent fallible assessment information falls directly on the achievement of individual learners and their academic and personal self-concept (Stiggins, Conklin, & Bridgeford, 1986). Yet others have reported that some teachers feel quite comfortable with their level of expertise in assessment practice and have no desire for further training (Gullickson, 1984; Schafer & Lissitz, 1987; Stiggins & Bridgeford, 1985). Stiggins (1991) attributed this comfort to a lack of understanding about what it takes to produce high-quality achievement data. Teachers who are "assessment illiterate," Stiggins proposes, are willing to accept achievement data at face value and lack the tools to be critical consumers of them. Many teachers, however, indicate concern about and interest in improving the quality of their classroom assessments (Stiggins et al., 1986), but inservice opportunities are rare (Stiggins, 1988) and the fundamentals of most measurement courses and textbooks are irrelevant to classroom teachers (McLean, 1985). Without concerted efforts to intervene this pervasive lack of assessment literacy is likely to continue because most teacher training (and administrator training) programs do not require courses in assessment and evaluation (Schafer & Lissitz, 1987). We need to develop further teachers' assessment literacy and proficiency if educational reform efforts are to be successful. And we need research-based knowledge to inform intervention strategies.

The research literature is replete with breaches in our knowledge about student evaluation practice. The purpose of the present investigation is to address some of these deficiencies. We wanted to know if, under reasonable circumstances, teachers could be trained to reliably assess student higher order thinking skills using an externally developed and validated test and student growth scheme. Because such activities are not common once students reach the senior elementary grades, we chose to focus on grade 7 and 8 teachers. We conducted two statistically independent studies of teachers' ability to assess students' correlational reasoning skills.

Experiment 1

The purpose of Experiment 1 was to compare novice and expert markers' assessment of correlational reasoning skills of students at various stages of development in their correlational proficiency. The study was a previously unreported component of a research program to develop valid assessment measures for complex skills. Data concerning substantive aspects of the research are reported elsewhere (Ross & Cousins, in press-a).

Method

Subjects

Subjects were a convenience sample of 20 students from each of grades 5, 7, 9, 11, and OAC (Ontario Academic Credit, graduating year) from four medium-sized school boards in east-central Ontario. Eight experienced grade 7 teachers from three of the same school systems volunteered to participate in an in-service program on correlational reasoning. Their participation in the research was part of the program.

Instruments

Two forms (swimming and medicine) of a correlational reasoning instrument were used in the present study. The instrument was designed to provide a complex task corresponding to correlational problems likely to be encountered by students in everyday contexts. The task for students was multivariate, included continuous as well as dichotomous variables, and placed limited knowledge demands on students.

In the swimming version subjects were given the age, swimming speed, and sex of 18 children and were asked to find out if swimming was related to age. The instructions were:

> Mrs. Campbell, the swimming coach, tests the swimming team every week. Here are the results from last Saturday.... Find out if there was a relationship between the height of students and how fast they swam. Use only the information in the table. Show your work.

Scoring was based on a four-dimensional student growth scheme, the development and validation of which is reported by Ross and Cousins (in press-a). The four dimensions and characteristics associated with them were:

- *Organizing*: extent to which students systematically organized and reordered the 18 data points; the scale ranged from 0, no productive display, to 4, data rearranged from lowest to highest on two variables and a control;
- *Locating*: the systematic selection of cases to decide whether a relationship existed; the scale ranged from 0, random or unsystematic selection of cases, to 2, enough cases selected to make a decision;
- *Synthesizing*: students' propensity to summarize the data; the scale ranged from 0, no attempt to summarize, to 3, trend line used to summarize the relationship; and
- *Concluding*: the extent to which students drew the correct three-variable conclusion; the scale ranged from 0, failure to describe any correct relationship, to 3, correct three-variable relationship described.

Only the end points for each of these scales is described. For each, intermediate levels of performance were made explicit. In the data provided, age and swimming speed were only weakly related until sex was controlled; age and

speed were strongly and positively correlated for girls, but no relationship existed for boys. The medicine version of the instrument was similar to that for swimming with one important distinction: data presented supported a counter-intuitive finding. The instrument asked students to determine whether a relationship existed between amount of vaccine administered to students and number of flu cases.

> One winter there was a lot of flu. All the students were given the same medicine to protect them from getting it. The amount of medicine given to the students was different in each school. Some students got the flu and some didn't.... Find out if there was a relationship between amount of medicine and number of cases of flu. Use only the information in the table. Show your work.

The data presented supported the conclusion that the number of flu cases increased with the amount of vaccine given in pill form; in needle form there was no relationship between vaccine and flu cases.

The validity and reliability of the instruments were established by Ross and Cousins (in press-a; in press-b). In previous administrations intercoder agreement was found to exceed .90 and intracoder agreement after a two-week interval was perfect (Ross & Cousins, in press-a). Also, Ross and Cousins (in press-a) reported that student performance was found to improve linearly with grade level and depended on test form; the swimming version was found to be easier than the medicine one. In the present study responses were coded by a trained tester (expert) and teachers (novices) using the explicit four skill growth scheme described above.

Treatment

The two forms of the instrument were randomly assigned so that each was completed by 10 students in each grade. (Each student completed one test on one occasion.) The responses were randomly assigned to 10 packages of vaccine items and 10 packages of swimming items. Each package contained one response from each of the five grades, arranged in random sequence. Eight vaccine and eight swimming packages were selected for use in the present study (i.e., $N = 80$). In the training workshop teachers were provided with an instructional module and participated in activities that illustrated what correlational reasoning is and how to teach it. Teachers were given copies of the growth scheme and were shown how to apply it in assessing correlational reasoning. Opportunities for practice and feedback were provided. Each teacher was randomly assigned two of the 16 packages in a counterbalanced design such that each package was marked independently by two teachers. Subsequently, an expert marker independently evaluated each of the 80 student responses using the same growth scheme.

Design and Analysis

The design was a 5 x 2 x 3 factorial multivariate analysis of variance with repeated measures. We used SPSSpc to analyze the data. The dependent variables were scores on organizing, locating, synthesizing, and concluding skill dimensions. The between group factors were grade (5 levels: 5, 7, 9, 11, and OAC) and test form (2 levels: swimming and medicine). The repeated measures factor was marker (3 levels: novice 1, novice 2, and expert). Each of the novice conditions contained composite marks from all eight teachers in the sample. Each student response, then, was independently marked by two teach-

ers and the expert. We used orthonormalized contrasts to compare the combined novice scores with those from the expert.

Our stringent criterion for marker accuracy was proportion of exact agreement with the expert marker. We also compared novice marker conditions with each other. As a less stringent criterion we examined Spearman rank order correlations between novice marker conditions and the expert for each of the four skill dimensions.

Hypotheses

Three hypotheses were evaluated in Experiment 1

Hypothesis 1 was the null hypothesis that no differences would be observed over novice and expert marker conditions. In other words, novice markers will meet expert standards in grading each of the skill dimensions. We predicted that training in how to teach correlational reasoning would enable teachers to apply the growth scheme with a reasonable degree of accuracy as reflected by high proportions of exact agreement and high positive correlations with the expert standard. Although we did not expect accuracy to be as high as intra- and intercoder data reported by Ross and Cousins (in press-a), coefficients were expected to exceed .80.

Hypothesis 2 predicted that student performance would increase linearly by grade for each of the skill dimensions. Ross and Cousins (in press-a) found such a relationship using the expert marker only. We reasoned that novice markers would be sensitive to increases in student growth and that the linear main effect for grade would be preserved.

Hypothesis 3. Student performance on the swimming instrument would exceed that for the medicine test form for each of the skill dimensions according to Hypothesis 3. Again, we predicted that novice markers would be sensitive to differences in student performance attributable to test form difficulty (Ross & Cousins, in press-a).

No statistical interactions among factors were predicted.

Results and Discussion

Means and standard deviations for the four skill dimensions by each cell in the 5 x 2 x 3 matrix appear in Table 1. Because the student sample was uninstructed in correlational reasoning, several responses were graded as zero, especially for younger students.

Table 2 and Table 3 show proportion of exact agreement data and correlations between novice and expert scores, respectively. The more stringent criterion, proportion of exact agreement with the expert, revealed that novice markers were unable to reach a high level of proficiency in their application of the growth scheme. Table 2 suggests that there was higher consistency between novice raters than for expert-novice comparisons and that agreement with the expert was highest for organizing and synthesizing skills. Marking the locating and concluding dimensions was problematic. Table 2 also reveals that novice markers were consistently more conservative on skill dimensions where there was less agreement with the expert (i.e., locating and concluding). Data reported in Table 3 divulged similar conclusions. Correlations between novice and expert marks were positive and small, albeit statistically significant in the main.

Table 1
Means (and Standard Deviations) for Skill Dimensions by Marker, Grade Level, and Test Form (Experiment 1, N = 80)

Skill Dimension	Grade	Form[a]	N	Marker: Novice 1		Novice 2		Expert	
	5	S	8	0.25	(0.46)	0.37	(0.74)	0.00	(0.00)
		M	8	0.13	(0.35)	0.13	(0.35)	0.00	(0.00)
	7	S	8	0.50	(0.75)	0.13	(0.35)	0.10	(0.32)
		M	8	0.37	(0.52)	1.00	(1.19)	0.00	(0.00)
Organizing	9	S	8	1.00	(1.31)	0.63	(1.07)	0.10	(0.32)
(0-4)		M	8	0.13	(0.35)	0.13	(0.34)	0.10	(0.32)
	11	S	8	0.50	(1.07)	0.25	(0.46)	1.00	(1.16)
		M	8	0.38	(1.07)	0.00	(0.00)	0.60	(0.97)
	OAC	S	8	2.13	(1.55)	2.00	(0.93)	1.20	(1.40)
		M	8	0.13	(0.35)	0.50	(1.07)	0.40	(0.97)
	5	S	8	0.50	(0.76)	0.63	(0.74)	0.20	(0.63)
		M	8	0.50	(0.76)	0.38	(0.74)	0.00	(0.00)
	7	S	8	0.63	(0.92)	0.63	(0.91)	1.10	(0.74)
		M	8	1.13	(0.99)	0.88	(0.99)	0.00	(0.00)
Locating	9	S	8	1.00	(0.93)	0.88	(0.83)	0.60	(0.84)
(0-2)		M	8	0.25	(0.71)	0.50	(0.93)	0.70	(0.95)
	11	S	8	0.63	(0.91)	0.50	(0.93)	1.80	(0.42)
		M	8	0.25	(0.46)	0.38	(0.52)	1.70	(0.48)
	OAC	S	8	1.63	(0.74)	1.88	(0.35)	1.50	(0.71)
		M	8	0.63	(0.74)	0.63	(0.92)	1.10	(0.74)
	5	S	8	0.25	(0.46)	0.25	(0.46)	0.00	(0.00)
		M	8	0.13	(0.35)	0.25	(0.46)	0.00	(0.00)
	7	S	8	0.13	(0.35)	0.50	(0.53)	0.10	(0.32)
		M	8	0.38	(0.51)	0.63	(0.74)	0.00	(0.00)
Synthesizing	9	S	8	0.25	(0.46)	0.25	(0.46)	0.10	(0.32)
(0-3)		M	8	0.13	(0.35)	0.25	(0.46)	0.00	(0.00)
	11	S	8	0.13	(0.35)	0.25	(0.46)	0.20	(0.42)
		M	8	0.00	(0.00)	0.25	(0.46)	0.20	(0.42)
	OAC	S	8	0.75	(0.71)	0.86	(0.35)	0.40	(0.97)
		M	8	0.25	(0.46)	0.38	(0.52)	0.10	(0.31)
	5	S	8	0.13	(0.35)	0.00	(0.00)	0.40	(0.84)
		M	8	0.50	(0.76)	0.75	(0.71)	0.00	(0.00)
	7	S	8	0.38	(0.74)	0.25	(0.46)	0.90	(0.88)
		M	8	0.50	(0.76)	0.38	(0.52)	0.30	(0.48)
Concluding	9	S	8	0.63	(0.91)	0.13	(0.35)	1.10	(0.88)
(0-3)		M	8	0.00	(0.00)	0.13	(0.35)	0.20	(0.42)
	11	S	8	0.38	(0.74)	0.75	(1.04)	1.20	(1.03)
		M	8	0.00	(0.00)	0.13	(0.36)	1.10	(0.74)
	OAC	S	8	1.13	(1.64)	1.63	(0.92)	2.10	(0.57)
		M	8	0.63	(1.41)	0.63	(0.74)	1.40	(0.97)

[a]S=Swimming. M=Medicine.

Table 2
Proportion Agreement with Expert for Skill Dimensions
(Experiment 1, N = 80)

	Comparison						
	Novice 1 vs. Novice 2		*Novice 1[a] vs. Expert*		*Novice 2[a] vs. Expert*		
Skill Dimension	*Exact Agreement*	*Below Expert*	*Exact Agreement*	*Above Expert*	*Below Expert*	*Exact Agreement*	*Above Expert*
Organizing	.73	.10	.64	.26	.09	.70	.21
Locating	.80	.31	.46	.23	.29	.51	.20
Synthesizing	.78	.06	.75	.19	.04	.63	.33
Concluding	.58	.41	.44	.15	.37	.50	.13

[a]Proportions sum to 1.0 within comparison categories.

The results indicate that support for the null hypothesis (Hypothesis 1) was not forthcoming. Clearly teachers were unable to meet the standard set by the expert marker. This finding raises questions about the efficacy of the training program. Teachers who participated were uninstructed in correlational reasoning prior to the half-day session. Conjecture that the training was insufficient for teachers to internalize the important dimensions of correlational reasoning and to translate this into assessment practice in such a short time span may be sustained. Although we provided opportunities for teachers to practice applying the growth scheme to examples of student performance, it may have been unreasonable to expect them to have sufficiently mastered the material without applying it in a more meaningful context such as their own classroom.

Table 4 provides further evidence of nonsupport for the null hypothesis. A statistically significant multivariate main effect for marker was obtained. Orthonormalized univariate contrasts revealed that the combined novice mean scores differed from those of the expert for each of the organizing, $F(1,70)=11.66$, $p<.001$, synthesizing, $F(1,70)=8.26$, $p<.001$, and concluding,

Table 3
Spearman Rank Order Correlation of Novice vs Expert Marks for Skill Dimensions (Experiment 1, N = 80)

	Correlations	
Skill Dimension	*Novice 1*	*Novice 2*
Organizing	.28**	.29**
Locating	.05	.22*
Synthesizing	.18*	.12
Concluding	.18*	.33***

*$p<.05$.**$p. <01$. ***$p.<001$.

Table 4
Sources of Variation for 5x2x3 Multivariate Analysis of Variance with Repeated Measures (Experiment 1, N = 80)

Source	*Univariate df*[a]	*Hotelling's Approx. F*	*Multivariate df*	*Probability*
Grade (G)	4	2.86	16, 262	<.001
Form (F)	1	2.42	4, 67	=0.56[b]
G x F	4	2.01	16, 262	<.01
Error A	70			
Marker (M)	1	9.20	8, 63	<.001
M x G	4	2.31	32, 246	<.001
M x F	1	1.27	8, 63	n.s.
M x G x F	4	1.47	32, 246	=.057[b]
Error B	70			
Total	159			

[a]Assumes orthonormalized contrasts of novice 1 and 2 vs. expert marker.
[b]Arguably significant difference.

$F(1,70)=24.09$, $p<.001$, skill dimensions. Novice mean scores exceeded expert scores for organizing, $M=0.57$ vs. 0.35, and synthesizing, $M=0.31$ vs. 0.11, dimensions, but were lower than the expert's for concluding, $M=0.46$ vs. 0.87. There were no differences for the locating skill dimension. This pattern agrees with the data in Table 2.

Table 4 reveals a statistically significant multivariate main effect for grade. Univariate linear contrasts divulged that a statistically significant linear component was discernible for each of the organizing, $t(70)= 4.08$, $p<.001$, locating, $t(70)=4.28$, $p<.001$, synthesizing, $t(70)=2.06$, $p<.05$, and concluding, $t(70)=4.86$, $p<.001$, dimensions. These data agree with the findings from Ross and Cousins (in press-a) and provide support for Hypothesis 2. After instruction teachers were able to differentiate variation in student performance in correlational reasoning attributable to age.

A multivariate main effect for test form is also apparent in Table 4. The swimming form was found to be easier than the medicine version for each of the organizing, $F(1,70)=9.09$, $p<.01$, locating, $F(1,70)=6.11$, $p<.05$, and concluding, $F(1,70)=4.86$, $p<.05$, skill dimensions, but was not different for the synthesizing skill. The means for swimming versus medicine versions for organizing, locating, and concluding were $M=0.65$ vs. 0.25, $M=0.93$ vs. 0.61, and $M=0.29$ vs. 0.20 respectively. These data agree with those reported by Ross and Cousins (in press-a) and provide support for Hypothesis 3. Teachers exposed to the treatment were able to detect variation in student performance attributable to test difficulty.

Several unexpected multivariate interaction effects are apparent in Table 4 and invite inspection. A grade by test form interaction was found for the organizing skill only, $F(4,70)=5.27$, $p<.001$. Inspection of the means in Table 1 reveals that the gap in scores favoring the swimming version of the instrument was greatly exacerbated for the OAC subsample. Second, a marker by grade

interaction was found to be statistically significant in the multivariate case (see Table 4). The univariate tests of this interaction divulged significant orthonormalized contrast effects comparing novice and expert marks for each of the organizing, $F(4,70)=2.48$, $p<.05$, locating, $F(4,70)=9.53$, $p<.001$, and concluding, $F(4,70)=5.83$, $p<.001$, skill dimensions. Inspection of Table 1 reveals that novices tended to be more generous with younger student responses, particularly grades 5 and 7, but much more stringent with those from grade 11. Finally, there was evidence of a three-way, marker by grade by test form multivariate interaction that was reflected in univariate tests of novice-expert orthonormalized contrasts for locating, $F(4,70)=4.80$, $p<.01$, and concluding, $F(4,70)=2.59$, $p<.05$. Inspection of the means suggests that the pattern of more generous grades awarded to younger students by novice markers yielding to more conservative marks in the higher grades was more pronounced for the more difficult medicine version of the instrument.

The two- and three-way interactions were not predicted and are difficult to interpret. While undertaking the task, markers were not explicitly aware of respondents' grade level. Perhaps teachers were more sensitive to level of development through response cues such as printing style, grammar, and diagrammatic representation, and were compelled to show empathy toward those at earlier stages of development. In short, teachers may have shown a greater predilection to make inferences for these respondents. Conversely, they may have been more demanding for students who were evidently from higher grades operating on the belief that these students should be much more adept at articulating clearly their responses. This interpretation suggests that teachers were influenced by stimuli other than levels on the growth scheme in assigning marks. Similar evidence regarding evaluative criteria has been cited by others (e.g., Natriello, 1987; Stiggins et al., 1989; Wilson, 1990). If this is true, the extent to which teachers had internalized the growth scheme would be a pivotal variable in predicting marking accuracy.

Experiment 2

The purpose of Experiment 2 was to investigate whether increased understanding of correlational reasoning concepts would enhance teachers' proficiency in evaluating the higher order skills. In addition, we wanted to make the context for assessment more ecologically valid. Teachers normally make judgments about students from a single or split grade; our multigrade sample in Experiment 1 may have been viewed as an artificial situation. Furthermore, teachers normally mark their own students' responses. We gave them an opportunity to do so in Experiment 2.

Method

Subjects

Subjects were 40 grade 7 and 8 teachers from a large metropolitan suburban public school board in central Ontario. The teachers volunteered for an inservice program on teaching correlational reasoning. Most teachers taught geography but a few specialized in other subject areas such as math and English. Thirty-one teachers attended the second of two sessions of the inservice and participated in the study. The grade 7 and 8 students who were members of classes of participating teachers also took part in the research. The responses of

10 students were randomly selected by the researchers from each class. Three of these students were dropped because they provided incomplete data, which reduced the total to 307.

Instruments

The swimming and medicine forms of the correlational reasoning instrument employed in Experiment 1 were also used in the current study. Procedures for scoring were identical to those described under that study.

Markers were provided with mark recording forms with the selected students' names listed. Finally, a self-report checklist was provided for teachers to indicate the extent to which they had covered the instructional module in their class. The checks were summed to yield scores ranging from 0 to 16. This instrument has been used in our previous research (Cousins & Ross, in press; Ross & Cousins, 1992) and it was found to predict achievement in one of our studies (Ross & Cousins, in press-b).

Treatment

Teachers attended two half-day sessions on correlational reasoning. In session 1 they were introduced to concepts and teaching methods and provided with an instructional module and a set of instructions for data collection. The workshop included opportunities to work on practice exercises and obtain feedback.

Teachers then returned to their classrooms and over a two and one half month period implemented the instructional unit. Teachers were instructed to collect pretest data immediately following session 1 and to submit the unmarked pretest responses to the researchers. They then proceeded to implement the unit at their convenience and to administer the posttest and bring responses with them to session 2. We recommended implementing the module in 10 40-minute periods. The instructional module and implementation sequence are described elsewhere (Ross & Cousins, in press-a, 1992).

Session 2 focused on the assessment of student performance in correlational reasoning. The growth scheme was reviewed, teachers practiced marking using data from other studies and feedback on the consistency of their appraisals was given. Each teacher then individually marked the pre- and posttests of 10 students randomly selected from his or her class. Two teachers failed to complete the unit and did not administer the posttest, and two additional teachers did not attend session 1 and therefore did not implement the unit. These teachers were all provided with student pre- and post- data from other teachers' classes for the session 2 workshop.

Subsequently, the expert marker from Experiment 1 independently graded all 307 responses.

Design and Analysis

Data were aggregated to the class level (N=31) and SPSSpc procedures were used for all analyses. The design was a 3 x 2 x 2 factorial multivariate analysis of variance with repeated measures. The dependent variables were scores on the four skill dimensions from the growth scheme. The between group factor was implementation (3 levels: high, moderate, low). The implementation variable was norm-based and constructed by dividing the sample into thirds. Ns were 11, 9, and 11 for high, moderate, and low groups respectively. Implementation checklist sums ranged from 0 to 16, and on average teachers had stu-

dents do over 10 of the optional exercises and quizzes. In most cases, teachers were able to cover each of the five lessons. The cut points for the independent variable were 10 or below for the low group, 11 to 13 for the moderate group, and 14 or greater for the high group. We used polynomial orthogonal contrasts to test for a linear trend main effect for implementation.

The within group factors were marker (2 levels: novice, expert) and test interval (2 levels: pretest and posttest). As in Experiment 1 the novice condition was a composite of marks from all teachers in the sample, whereas the expert marked all responses.

Both the stringent proportion of exact agreement with the expert and more lenient Spearman rank order correlations between novice markers and the expert were examined. These analyses were broken down by the four skill dimensions.

Hypotheses

Four hypotheses were tested in the current study.

Hypothesis 1. As in Experiment 1, Hypothesis 1 was the null hypothesis that mean novice grades would not differ from those of the expert. We predicted that value-added instruction in correlational reasoning would enable teachers to apply the growth scheme with a high degree of accuracy. The added value was reflected in their implementation of the unit in their own classrooms including the collection of pre- and posttest data from their own students. We assumed that such treatment would significantly improve their understanding of correlational reasoning and concomitantly their assessment of student performance on such tasks. Again, we predicted that accuracy levels would exceed .80 on the agreement and correlation standards and that mean scores would not vary over markers as reflected by the absence of a statistically significant marker main effect in the analysis of variance.

Hypothesis 2 was that marker accuracy would increase with implementation intensity. We predicted that more intense implementation of the unit (i.e., greater coverage) would deepen teachers' understanding of correlational reasoning concepts and that their assessment of student performance would be more likely to meet the standard set by the expert. Statistically, then, we did not expect an implementation main effect, but we did predict a significant two-way interaction for marker by implementation. That is, novice and expert marks would be more likely to converge for markers with greater implementation experience.

Hypothesis 3. Posttest performance, according to Hypothesis 3, would be graded higher than pretest performance. Our previous research has demonstrated repeatedly that the instructional treatment is very powerful in improving students' understanding and achievement in correlational reasoning (Cousins & Ross, in press; Ross & Cousins, in press-a, in press-b, 1992). Given the value-added treatment, teachers were expected to be sensitive to such demonstration of student growth in performance as reflected by a main effect for treatment interval favoring posttest over pretest means.

Hypothesis 4. A fourth hypothesis predicted that pre- and posttest scores would diverge as implementation increased. We predicted that more intense implementation of the instructional module would lead to superior posttest scores and that markers would be sensitive to such evidence of student

Table 5
Means (and Standard Deviations) for Skill Dimensions by Marker and Test Interval (Experiment 2, N = 31)

			Marker			
			Novice		Expert	
			Test Interval			
Skill Dimension	Implementation	N	Pre	Post	Pre	Post
Organizing	Low	11	0.90 (0.64)	2.32 (0.10)	0.69 (0.69)	2.47 (0.72)
(0-4)	Mod	9	1.01 (1.11)	2.53 (0.76)	0.72 (1.13)	2.58 (0.86)
	High	11	0.84 (0.65)	3.35 (0.79)	0.35 (0.63)	3.22 (0.87)
Locating	Low	11	1.03 (0.57)	1.78 (0.60)	1.36 (0.32)	1.96 (0.07)
(0-2)	Mod	9	1.13 (0.69)	2.00 (0.00)	1.13 (0.73)	1.99 (0.03)
	High	11	0.97 (0.62)	1.98 (0.06)	1.29 (0.57)	2.00 (0.00)
Synthesizing	Low	11	0.64 (0.52)	2.27 (0.91)	0.40 (0.49)	2.67 (0.33)
(0-3)	Mod	9	0.62 (0.79)	2.68 (0.37)	0.32 (0.65)	2.46 (0.95)
	High	11	0.46 (0.48)	2.74 (0.44)	0.28 (0.59)	2.84 (0.30)
Concluding	Low	11	0.64 (0.49)	0.89 (0.44)	0.81 (0.36)	0.97 (0.42)
(0-3)	Mod	9	0.37 (0.28)	1.06 (0.62)	0.75 (0.35)	1.37 (0.54)
	High	11	0.60 (0.31)	1.64 (0.59)	0.85 (0.32)	1.32 (0.71)

growth. This hypothesis would be confirmed by a statistically significant test interval by implementation interaction.

No other statistically significant interactions were predicted.

Results and Discussion

Means and standard deviations for the four skill dimensions by each cell of the 3 x 2 x 2 matrix appear in Table 5. The floor effect for results from uninstructed students (pretest) apparent in Experiment 1 (see Table 1) was not as evident here, probably because means were based on 10 students.

Tables 6 and 7 show data pertinent to Hypothesis 1. Proportion of exact agreement data are reported in Table 6. Although some improvement over Experiment 1 findings in marker accuracy was evident (compare Table 2) agreement remained lower than expected. Generally, posttest grading was more accurate (perhaps because higher levels of performance were evident) and marking on locating and synthesizing dimensions was respectable. Little improvement was evident for the organizing and concluding dimensions, with the latter continuing to create discrepancies between novice and expert markers. With the exception of organizing there was a modest tendency for novice markers to be more generous than the expert on posttest performance. For pretests, this was the case for locating and concluding, but the reverse was true for the other two dimensions. This pattern is at odds with observations from Experiment 1. Posttest grades may have been inflated by teachers who believed they had successfully implemented the unit. Correlation data reported in Table 7 were modestly encouraging. Almost all correlations were positive and statistically greater than zero, but none reached .80 suggesting

Table 6
Proportion Agreement with Expert for Skill Dimensions, by Test Interval (Experiment 2, N = 31)[a]

Skill Dimension	*Below Expert*		*Exact Agreement*		*Above Expert*	
	Pre	*Post*	*Pre*	*Post*	*Pre*	*Post*
Organizing	.30	.13	.64	.76	.06	.11
Locating	.05	.01	.76	.95	.20	.04
Synthesizing	.27	.06	.66	.82	.06	.12
Concluding	.14	.27	.56	.46	.30	.27

[a]Proportions add to 1.0 within pre- or posttest intervals for each skill dimension.

that covariation between expert and novice scores was not high. Noticeable improvement over Experiment 1 results (compare Table 3) was evident, however.

Marginal support for the null hypothesis (Hypothesis 1) was apparent, but results from the multivariate factorial analysis of variance establish further limits on confidence in the success of the marking treatment. Table 8 shows a statistically significant multivariate main effect for marker. However, univariate tests divulged that differences were observed only for the locating skill dimension, $F(1,28)=5.25$, $p<.05$, although concluding was approaching significance, $F(1,28)=3.84$, $p=.06$, with alpha set at .05. For locating and concluding skills experts gave more generous marks than novices: $M=1.63$ vs 1.48, and $M=1.01$ vs 0.88, respectively. This finding agrees with Experiment 1 results.

Improvement over Experiment 1 results regarding Hypothesis 1 was evident, but Experiment 2 findings were less than satisfactory. It is likely that greater experience with correlational reasoning either through more intensive inservice or repeated classroom application would improve teachers' facility in assessing the higher order skills. Further support for this assertion comes from results pertinent to Hypothesis 2. As predicted, a main effect for implementation was not found, but a statistically significant two-way marker by im-

Table 7
Spearman Rank Order Correlation of Novice vs. Expert Marks for Skill Dimensions by Test Intervals (Experiment 2)

	Student Level		*Class Level*	
	Pre (N = 307)	*Post (N = 304)*	*Pre (N = 31)*	*Post (N = 31)*
Organizing	.59**	.63**	.72**	.65**
Locating	.68**	.32**	.36*	.42**
Synthesizing	.41**	.78**	.64**	.78**
Concuding	.38**	.36**	.65**	.18

*$p<.05$. **$p<.001$.

Table 8
Sources of Variation for 3x2x2 Multivariate Analysis of Variance with Repeated Measures (Experiment 2, N = 31)

Source	*Univariate df*	*Hotelling's Approx. F*	*Multivariate df*	*Probability*
Implementation (I)	2	0.61	8, 48	n.s.
Error A	28			
Marker (M)	1	7.04	4, 25	<.001
MxI	2	2.95	8, 48	<.01
Error B	28			
Test Interval (T)	1	56.83	4, 25	<.001
TxI	2	2.35	8, 48	<.05
Error C	28			
MxT	1	10.83	4, 25	<.001
MxTxI	2	2.19	8, 48	<.05
Error D	28			
Total	123			

plementation multivariate interaction emerged (see Table 8). None of the univariate tests of the interaction were statistically significant, however. The closest skill to achieving significance was concluding, $F(2,28)=2.04$, $p=.148$, and indeed marginal support of the Hypothesis 2 was evident with novice and expert marks converging at the highest level of implementation. A linear trend was not palpable, however. These findings do not sufficiently support the conclusion that teachers' marks converged with those of the expert as implementation intensity increased. An explanation for this unanticipated finding may partially reside in the fact that the implementation checklist was a self-report device. Although more valid indicators of implementation may improve the quality of measurement, findings concerning the predicted statistically significant interaction between test interval and implementation (reported below) were as expected, thereby providing at least partial support for the validity of the checklist. Questions concerning the relationship of this variable to marker proficiency remain to be answered.

Table 8 shows a statistically significant multivariate main effect for test interval. Hypothesis 3 was supported in every instance as revealed by univariate tests of organizing, $F(1,28)=94.79$, $p<.001$, locating, $F(1,28)=63.18$, $p<.001$, synthesizing, $F(1,28)=235.56$, $p<.001$ and concluding, $F(1,28)=62.74$, $p<.001$, skill dimensions. Pretest versus posttest means were $M=0.75$ vs. 2.76, $M=1.15$ vs. 1.95, $M=0.46$ vs. 2.61, and $M=0.68$ vs. 1.21 respectively, for the four skills. These findings attest to teachers' ability to observe and accordingly assess growth in student performance.

A statistically significant multivariate interaction term for test interval by implementation surfaced from the analysis, thereby supporting Hypothesis 4. Univariate analyses pinpointed the differences for two of the skill dimensions: concluding, $F(2,28)=6.37$, $p<.01$, was clearly significant whereas organizing,

$F(2,28)=3.06$, $p=.063$, was only approaching significance with alpha set at .05. Close inspection of means in Table 5 reveals that indeed posttest scores were substantially better than pretest scores at higher levels of implementation. Significant differences were not found for locating or synthesizing skills.

Finally, one two-way and one three-way interaction turned out to be statistically significant. Although these remain difficult to interpret, they invite comment. First, a test interval by marker interaction was not expected; pre- and posttest performance differentials should have remained constant regardless of marker. In fact, the interaction was statistically significant for the organizing, $F(1,28)=6.08$, $p<.05$, locating, $F(2,28)=6.91$, $p<.01$, and synthesizing, $F(2,28)=15.66$, $p<.001$ skill dimensions and approached significance for concluding, $F(2,28)=3.81$, $p=.061$, with alpha set at .05. Inspection of the means in Table 5 suggests that although the expert was consistently slightly more generous than novices in awarding posttest grades, novices were more variable on pretest marks. In the case of organizing and synthesizing they were more generous than the expert, while at the same time being more conservative regarding locating and concluding pretest marks. One possible explanation is that teachers may have been more generous in situations where students have not had an opportunity to learn (i.e., pretest interval, more difficult skill) and less generous when students "should know it." The three-way interaction was restricted to the synthesizing skill, $F(2,28)=3.39$, $p<.05$. Inspection of the means suggested that the expectation that novice and expert mean grades would converge at higher levels of implementation was palpable for the post-, but not pretest interval. Possibly this finding is related to the fact that the range of student performance was higher on the posttest, although no theoretical justification is readily apparent.

Discussion

We sought to examine whether teachers' proficiency in evaluating higher order thinking skills could be enhanced through a reasonable amount of training. Experiment 1 yielded results that showed a discrepancy between teacher and expert assigned marks. However, findings revealed that although teachers were unable to meet the standard set, they were sensitive to variation in performance attributable to the developmental stage of the respondents and to test difficulty even in an assessment setting with limited ecological validity. In Experiment 2 we wanted to see if a training activity involving application of correlational reasoning instructional activities in the classroom would assist teachers in internalizing key concepts and principles and help them to achieve higher levels of accuracy in assessing student performance. Some improvement over one-shot instruction was evident, but the results were not dramatic. Teacher assigned grades were found to covary with those of the expert slightly better than in Experiment 1, but differences remained substantial. There was only modest evidence to suggest that grading proficiency depended on implementation intensity. Teachers were, however, sensitive to growth in student performance.

Because our goal is to improve classroom-based student assessment, it is appropriate to evaluate our results in terms of the purposes teachers have for student evaluation. Is it really necessary for teachers to reach the level of proficiency attained by our expert marker? One purpose of student assessment

is to make decisions about the quality of instruction (Crooks, 1988; Dorr-Bremme, 1983; Stiggins et al., 1989). In Experiment 2, teachers were able to detect the effects of instruction and to recognize growth in student performance suggesting that their level of proficiency in marking was sufficient for this purpose. Also, it is possible that at least some teachers had, and were evaluating with, a more generalized view of student performance that was sensitive to serendipitous learning and that was not restricted to a narrow set of objectives for students. This consideration may account for at least some of the observed discrepancies.

A second purpose is to make decisions about student progress, a need that precipitates much of the evaluation activity in schools (Wilson, 1989). In addition to being able to detect the effects of instruction (Experiment 2), teachers were also able to correctly distinguish student performance on the basis of age, confirming that their attained level of proficiency was adequate for this second purpose. However, the discrepancy between novice and expert marks raises the issue of fairness (Linn et al., 1991). The issue hinges on what other considerations entered into teachers' decisions in marking the correlational items. Technically, it would be better to separate these criteria from the overt criteria of the test. On the other hand, if other criteria are valid, it may be reasonable to combine them with teacher judgments about correlational reasoning on the grounds of efficiency. Others have shown that final grades are the consequence of mixing formative and summative data (Rogers, 1991) and that teachers tend to predict patterns of results for students from prior knowledge of them (Wilson, 1990). In order to fully appreciate whether the inclusion of other criteria in the correlational reasoning decisions was unfair to individuals, researchers would have to sample the range of decisions in search of systematic bias.

A third purpose of assessment is diagnosis of individual student performance with a goal of informing group placement or designing individualized instructional activities (Earl & Cousins, 1993; Hargreaves & Earl, 1990; Natriello, 1987). Evidence concerning teachers' proficiency in detecting variation due to age and instruction suggests that teachers would be able to appropriately engage in diagnostic activities. One use of diagnostic information would be to decide on placement in ability groups. The discrepancy between novice and expert marks suggests that the accuracy of such placement may be dubious. On the other hand, the imprecision may be due to variability in scores between tasks. For example, we found that medicine and swimming items produced different estimates of student performance. Other researchers report that inter-item variability is a major problem in the authentic assessment of performance on cognitively complex tasks (Linn et al., 1991). The issue is equally important if teachers are being urged to opt for equivalent mixed-ability grouping where procedures require knowledge of student performance (e.g., Slavin, 1990).

Several issues remain outstanding and form an agenda for further research. First, what level of implementation intensity constitutes sufficient intensity in order for teachers to successfully internalize correlational reasoning? For example, would teaching the unit for a second time have contributed to assessment skill development? Second, a possible confounding variable in Experiment 2 was knowledge of the student. We had teachers mark their own

students' work, although the expert remained blind as to student identity. We know from other reports that teachers sometimes formulate global impressions of students and assessments of their performance can be influenced by such impressions (Stiggins et al., 1989; Wilson, 1990). It would be interesting to compare teachers' assessment accuracy with samples of work from familiar and unfamiliar students. In this way we would be able to detect the extent to which assessment depends on criteria over and above performance. Third, we assumed that implementation intensity equated with teachers' knowledge of correlational reasoning concepts and facility with them. Our data suggest that this assumption might have been somewhat heroic. It would be useful to assess teachers' proficiency in understanding correlational reasoning as a covariate. The estimation of the effect of implementation intensity would be more precise under these circumstances. Fourth, despite the fact that senior elementary grades have been suggested to be more like secondary schools than secondary schools (Hargreaves & Earl, 1990), it seems likely that teachers of these grades would have more opportunity for exposure to alternative student assessment techniques than their secondary counterparts, by virtue of proximity to junior grades if nothing else. Two possibilities for future research are suggested: (a) replicate the findings in the secondary panel, and (b) measure and control for adeptness with performance assessment and other alternative evaluative techniques. Finally, although Experiment 2 provided a more ecologically valid context for assessment, the treatment may not have provided enough support for novice markers. We know from prior research that deeper levels of teachers' joint work will help to foster teachers' conceptual development (Cousins, Ross, & Maynes, in press; Little, 1990). Another viable direction for research would be to evaluate the impact of alternative treatments where teachers collaborate in reciprocal coaching, team teaching, or concurrent implementation arrangements. It seems likely that such training conditions would be more powerful in bringing about deeper understanding of assessment criteria and fostering the development of assessment skills.

The findings from the current study have significant implications for educational reform. Clearly the interconnections between curriculum and student evaluation are substantial, and it is not likely that fundamental change in one will occur with an absence of change in the other (Linn et al., 1991; Rogers, 1991). As we have seen, however, change in teachers' student assessment practice is not likely to be accomplished with great ease. Significant resources should be directed toward both pre- and inservice in this general domain, especially in training teachers to grade higher order thinking and cognitively complex skills. It is likely that teachers would have been more proficient at alternate assessment with tools developed by themselves. However, given the demands placed on teachers, it is not always plausible to expect them to develop with care valid and reliable instruments for evaluating thinking. Policy makers need to be cognizant of the interrelationships between evaluation and curriculum and to place much more emphasis on the former when significant change in classroom practice is the goal of reform.

Note

Funding for separate phases of the research reported here was provided by the Ontario Ministry of Education (Grant #81-1084) and the Social Sciences and Humanities Research Council of

Canada (Grant #410-90-0834). The opinions expressed within do not necessarily reflect those of the funding agencies. Anne Hogaboam-Gray and Peter Gray made valued contributions at various stages of the work.

References

Airasian, P.W. (1979, April). *The effects of standardized testing and test information on teachers' perceptions and practices*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Anderson, J.O. (1989). Evaluation of student achievement: Teacher practices and educational measurement. *Alberta Journal of Educational Research, 35*, 123-133.

Bateson, D.J. (1990). Measurement and evaluation practices of British Columbia science teachers. *Alberta Journal of Educational Research, 36*, 45-51.

Bloom, B.S. (1984). The search for methods of group instruction as effective as one-to-one tutoring. *Educational Leadership, 41*, 4-17.

Bongar, C. (1982). Dissonant feedback about achievement and teachers' expectations. *Alberta Journal of Educational Research, 28*, 277-287.

Cousins, J.B., & Ross, J.A. (in press). Improving higher order thinking skills teaching "with" the computer: A comparative study. *Journal of Research on Computing in Education.*

Cousins, J.B., Ross, J.A., & Maynes, F.J. (in press). The reported nature and consequences of teachers' joint work in three exemplary schools. *Elementary School Journal.*

Crooks, T.J. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research, 58*, 438-481.

Dorr-Bremme, D.W. (1983). Assessing students: Teachers' routine practices and reasoning. *Evaluation Comment, 6*(4), 1-11.

Earl, L., & Cousins, J.B. (1993). Enhancing classroom assessment. In A. Hargreaves, D. Gérin-Lajoie, & K. Leithwood (Eds.), *Exemplary practices in the Transition Years: A review of research and theory* (pp. 2-31). Toronto: Queen's Printer.

Fleming, M., & Chambers, B. (1983). Teacher-made tests: Windows on the classroom. In W. Hathaway (Ed.), *Testing in the schools: New directions for testing and measurement*. San Francisco, CA: Jossey Bass.

Gullickson, A.R. (1984). Teacher perspectives of their instructional use of tests. *Journal of Educational Research, 77*, 244-248.

Hargreaves, A., & Earl, L. (1990). *The rights of passage*. Toronto: Ontario Ministry of Education.

Herman, J., & Dorr-Bremme, D. (1983). Uses of testing in the schools: A national profile. In W. Hathaway (Ed.), *Testing in the schools. New directions for testing and measurement* (No. 19, pp. 7-17). San Francisco, CA: Jossey Bass.

Linn, R.L., Baker, E.L. & Dunbar, S.B. (1991). Complex, performance-based assessment: Expectations and validation criteria. *Educational Researcher, 20*(8), 15-21.

Little, J.W. (1990). The persistence of privacy: Autonomy and initiative in teachers' professional relations. *Teachers College Record, 91*, 509-535.

MacIver, D. (1989). *Effective practices and structures for middle grades education*. Charleston, WV: Policy & Planning Centre, Appalachia Educational Laboratory. (ERIC Document Reproduction Service No. 306 668)

McLean, L.D. (1985). *The craft of student evaluation Canada*. Toronto: Canadian Education Association.

McLean, L.D. (1990). Time to replace the classroom test with authentic measurement. *Alberta Journal of Educational Research, 36*, 78-84.

Nagy, P., & Moorehead, R. (1990). Administrative response to classroom testing data: A problem solving perspective. *Alberta Journal of Educational Research, 36*, 18-34.

Natriello, G. (1987). The impact of evaluation processes in students. *Educational Psychologist, 22*, 155-175.

Paris, S.G., Lawton, T.A., Turner, J.C., & Roth, J.L. (1991). A development perspective on standardized achievement testing. *Educational Researcher, 20*(5), 12-20.

Rogers, W.T. (1991). Educational assessment in Canada. *Alberta Journal of Educational Research, 37*, 179-192.

Ross, J.A., & Cousins, J.B. (1992). *Development of correlational reasoning: Effects of alternate instructional strategies*. Final report. Social Sciences and Humanities Research Council of Canada, Grant #410-90-0834.

Ross, J.A., & Cousins, J.B. (in press-a). Patterns of student growth in reasoning about correlational problems. *Journal of Educational Psychology.*

Ross, J.A., & Cousins, J.B. (in press-b). Enhancing secondary school students' acquisition of correlational reasoning skills. *Research on Science and Technological Education.*

Salmon-Cox, L. (1981). Teachers and standardized achievement tests: What's really happening? *Phi Delta Kappan, 62,* 631-634.

Schafer, W.D., & Lissitz, R.W. (1987). Measurement training for school personnel: Recommendations and reality. *Journal of Teacher Education, 38*(3), 57-63.

Shavelson, R.J., Baxter, G.P., & Pine, J. (1992). Performance assessments: Political rhetoric and measurement reality. *Educational Researcher, 21*(4), 22-27.

Shepard, L.A. (1991). Psychometricians' beliefs about learning. *Educational Researcher, 20*(7), 2-10.

Sizer, T.R. (1984). *Horace's compromise: The dilemma of the American high school.* Boston, MA: Houghton Mifflin.

Slavin, R.E. (1980). Effects of individual learning expectations on student achievement. *Journal of Educational Psychology, 72,* 520-524.

Slavin, R.E. (1990). *Cooperative learning: Theory, research, and practice.* Englewood Cliffs, NJ: Prentice Hall.

Smith, M.L. (1991). Put to the test: The effects of external testing on teachers. *Educational Researcher, 20,* 8-11.

Sproull, L., & Zubrow, D. (1981). Standardized testing from the administration perspective. *Phi Delta Kappan, 6,* 631-634.

Stiggins, R.J. (1987, Fall). Design and development of performance assessment. *Educational Measurement: Issues and Practices, 6*(3), 33-42.

Stiggins, R.J. (1988). Revitalizing classroom assessment: The highest instructional priority. *Phi Delta Kappan, 69,* 363-368.

Stiggins, R.J. (1991). Assessment literacy. *Phi Delta Kappan, 72,* 534-539.

Stiggins, R.J., & Bridgeford, N.J. (1985). The ecology of classroom assessment. *Journal of Educational Measurement 22*(4), 271-286.

Stiggins, R.J., Conklin N., & Bridgeford N.J. (1986). Classroom assessment: A key to educational excellence. *Educational Measurement: Issues and Practices, 5*(2), 5-17.

Stiggins, R.J., Frisbie, D.A., & Griswold, P.A. (1989). Inside high school grading practices: Building a research agenda. *Educational Measurement: Issues and Practices, 8*(2)5-14.

Wiggins, G. (1989, May). A true test: Toward a more authentic and equitable assessment, *Phi Delta Kappan, 70,* 703-713.

Wilson, R.J. (1989). Evaluating student achievement in an Ontario high school. *Alberta Journal of Educational Research, 35,* 134-144.

Wilson, R.J. (1990). Classroom processes in evaluating student achievement. *Alberta Journal of Educational Research, 36,* 4-17.

Wilson, R.J., Rees, R., & Connock, M. (1989, Winter). Classroom practices in evaluation: A progress report from the Ecology of Evaluation Project. *Research Forum,* 4-7.

The Alberta Journal of Educational Research *Vol. XXXIX, No. 3, September 1993, 319-334*

George H. Buck
University of Alberta

Faculty of Education, University of Alberta: A 30-year Gestation[1]

Although the first president of the University of Alberta desired the establishment of a faculty of education as early as 1912, it took 30 years before one was established. Several unsuccessful attempts were made by the administration of the University of Alberta and by particular politicians within the provincial government to create a faculty of education before its establishment in 1942. These are described and analyzed, along with factors that contributed to the failure of these attempts. The events occurring at the University of Alberta are placed within the context of the status and development of faculties of education at other state-funded universities in North America. The importance of considering antecedent problems and the approaches used to solve them successfully is noted.

Dès 1912, le premier président de l'Université de l'Alberta désirait l'établissement d'une faculté de pédagogie. Son souhait se réalisait 30 ans plus tard. Avant sa création en 1942, le corps administratif de l'Université de l'Alberta et certains politiciens particuliers du gouvernement provincial avaient tenté à plusieurs reprises, mais sans succès, d'y créer une faculté de pédagogie. Cette article décrit et analyse les facteurs qui ont contribués à ces nombreux échecs. Les événements qui se sont déroulés à l'Université de l'Alberta sont expliqués dans le contexte historique de la situation particulière du développement des facultés de pédagogie dans d'autres universités subventionnées par l'état en Amérique du Nord à l'époque. On note l'importance du rôle qu'ont joué les problèmes antérieures à la résolution des conflits et les méthodes qui ont réussies à les résoudre.

The festivities conducted in the University of Alberta during 1992 to commemorate the 50th anniversary of the formation of the Faculty of Education leave an impression that they are part of a birthday celebration. Although such celebration is not to be criticized, it does little to direct attention to antecedent events and actions that determined why and when the Faculty of Education was established. Although there are works describing aspects of the development of teacher education in Alberta (Chalmers, 1992; Coutts, 1979; Patterson, 1979), such accounts present either a general historical overview, or a consideration of particular aspects, the role of the normal schools for example, from phenomenological perspectives primarily. Little is presented on the identification, description, and analysis of factors within and without the University of Alberta that impinged on the development of teacher education programs there prior to 1942.

Although events of the past may be considered as interesting occurrences that should be regarded, in an antiquarian way, as being detached from the

---

George Buck is currently a lecturer in the Department of Educational Psychology. His interests include learning theory and its practical application to pedagogy, computers in education, aspects of social psychology, and the history and evolution of educational psychology.

present, it is important to bear in mind that what exists now is a function of the past. Moreover, it may be that some of the factors and conditions that affected the formation of the Faculty of Education may continue both to exist and to affect the present course of development. An exposition and analysis of the significant plans, events, developments, and factors will help explain why the process of establishing a faculty of education took 30 years to achieve. In light of the analysis provided of the unsuccessful and successful strategies employed in the years leading up to 1942, this study may prove useful to personnel in both Canadian and American universities faced with selecting strategies to contend successfully with current conditions that have the potential to halt and to reverse the continued development of faculties of education.

Initial Attempts

1912: The First Attempt

When the Province of Alberta was established in September 1905, responsibility for all education came under its purview. This extended to the training and certification of schoolteachers. One of the first steps taken by the new province in this regard was the establishment of a provincial normal school in Calgary (Alberta Department of Education, 1907). The normal school provided the training deemed necessary by the Department of Education for individuals desiring to become teachers and for existing teachers who either desired or who were compelled to upgrade their qualifications. By 1908 the Province established the University of Alberta to provide postsecondary education. Initially there was no connection between the single teacher training institution and the university. Although the normal school provided training for elementary and intermediate (junior high school) teachers to a reasonably satisfactory level, evidence indicates that some responsible individuals considered it to be inadequate both in the proper training of high school teachers and in conducting research in education (Chalmers, 1978; Coutts & Walker, 1964; Myrehaug, 1972).

By 1912, sensing this sentiment and combining it with his desire to create an innovative and trend-setting Canadian university, President Henry Marshall Tory (1864-1947) reported that a department of education in the university, "will be in full operation during the session 1912-1913" (Tory, 1912, p. 3).[2] Further to this new department, Tory states that it is, "in the interest of students who intend following the teaching profession though they [Education and Agriculture] both have a definite place in a course of study for a liberal education. They also bring the course of study into harmony with the new courses or study prescribed for the schools" (p. 3). Although he did not criticize the normal school directly, Tory did so through implication. By mentioning that education courses offered at the university would be a part of a *liberal education*, Tory showed he believed that programs offered by the university would be superior to those offered by the normal school.

Further evidence for this view is Tory's first appointee to teach education courses, the Reverend Dr. Samuel Dyde (1862-1947), an eminent educator of teachers who was not associated with the provincial Department of Education. Tory (1912) states, "Dr. S.W. Dyde, Principal of the Presbyterian College [in Edmonton, later named Robertson College] came to Alberta after a distinguished career both as student and teacher at Queen's University [Kingston,

Ontario]" (p. 6). Although the normal school employed a few people who had earned doctorates, most of the staff did not possess degrees above the baccalaureate (Coutts & Walker, 1964).[3]

Under Dyde's direction as a Special Lecturer in the Department of Philosophy and Psychology, the University of Alberta offered two senior courses in Education, the Philosophy of Education, for one hour a week, and the History of Education for two hours a week (University of Alberta, 1912, p. 71). In spite of these offerings, a department or a faculty of education was not established at this time.[4] Although President Tory headed the University of Alberta, matters of funding and decisions concerning the creation of departments and faculties were ultimately under the control of the provincial government through the auspices of the Board of Governors and the Senate.

The perception of need for additional facilities for teacher training was also shared by the provincial government by 1912. A decision had to be made whether the new facility would be a normal school or a faculty of education. G. Fred McNally, then a school inspector in the Department of Education, claimed that the provincial government decided to construct a normal school in the central Alberta town of Camrose, because of political pressure exerted by George Smith who represented the largely rural riding containing Camrose (Coutts & Walker, 1964). The erection of a large normal school so close to Edmonton provided the government with a reason to stop the university from creating a faculty of education. Moreover, McNally states that in an attempt to deflect the criticism that staff in normal schools possessed qualifications inferior to university professors, "Dr. James Collins Miller, fresh from Teachers College, Columbia University, was appointed principal" (Coutts & Walker, 1964, p. 45).[5]

Although further discussion about establishing a faculty of education at the University of Alberta ceased at this time, largely because of the actions of the provincial government, it was also the provincial government that would be responsible for the next attempt at establishing a faculty of education at the university.

1918: The Second Attempt

Early in 1918 the Ministers and Deputy Ministers of Education of British Columbia, Alberta, Saskatchewan, and Manitoba met in Calgary. At the meeting it was decided to augment the standard normal school term from four months to eight (letter, Kerr to Tory, May 15, 1918).[6] In Alberta, a ramification of this decision was that an additional teacher training institution was required, either another normal school or a faculty of education. Corresponding with President Tory, who was in England supervising the Khaki University,[7] Acting President Kerr (May 15, 1918) stated, "I don't think there is any question as to which course will be decided upon. It is almost sure that it will be the Faculty of Education ... Mr. Boyle [James R. Boyle, Minister of Education] has spoken casually of the possibility of getting going in the fall." Moreover, Kerr urged Tory to support the idea,

> It is a good plan to strike while the iron is hot and to encourage the establishment of the Faculty at the earliest possible moment ... Western Canada, as you know only too well, is the kind of country that presents problems overnight and if

> great opportunities are not to be lost, action, you will agree I think, must sometimes be taken promptly. (letter, Kerr to Tory, May 15, 1918)

Although in favor of establishing a faculty of education, Tory was not keen to have it started immediately, because he was concerned that a suitable qualified candidate for dean might not be available (letter, Tory to Kerr, June 25, 1918). Tory maintained that a dean of a faculty must possess a doctorate (letter, Tory to Kerr, August 13, 1918). In spite of Tory's reluctance, Kerr began to compile names of potential candidates. Kerr's quest to find suitable individuals who might be imported from other universities was not successful. Writing to Tory (May 15, 1918) Kerr states that he "heard of very few men that would be worth considering at all." The failure to find suitable external candidates was not caused entirely by a paucity of appropriate individuals with doctoral degrees. Tory (letter to Kerr, August 13, 1918) indicates that the person selected for dean must possess extensive knowledge of the process and structure of the educational system in Alberta. Efforts were then concentrated on finding a suitable candidate within the province.

Dyde was not considered, because he left the University of Alberta earlier in the year to assume the principalship of Presbyterian Theological College at Queen's University. Although Dr. Miller, the first Principal of the Camrose Normal School, possessed a doctorate, he was not liked by either Tory or Kerr, especially after some unpleasantness Miller caused while serving on several liaison committees between the University of Alberta and the Alberta Department of Education (letter, Kerr to Tory, May 17, 1916).[8] The problem of finding a suitable candidate for dean remained, however. Tory suggested that if the provincial government was determined to start a faculty of education in the fall of 1918, then G. Fred McNally would be considered satisfactory (letters: Tory to Kerr, June 25, 1918; Tory to Kerr, July 15, 1918).

McNally (1878-1965) had lived in Alberta since 1906 and had worked his way up in the Department of Education from school inspector to Principal of the Camrose Normal School (Coutts & Walker, 1964). McNally had also obtained an master's degree from the University of Alberta in 1911 (Coutts & Walker, 1964). Beginning in 1914 the Department of Education provided McNally with some released time to pursue a doctorate at Teachers College, Columbia University in the area of educational psychology, where he studied under Edward L. Thorndike (Coutts & Walker, 1964). At the time Kerr and Tory were searching for a candidate for dean it appeared as if McNally would complete the requirements for the PhD within a year (letter, Kerr to Tory, August 10, 1918).

Although both Tory and Kerr were satisfied with the idea of McNally as Dean of Education, Tory urged Kerr to persuade the provincial government to delay the creation of a faculty of education until the fall of 1919 (letter, Tory to Kerr, June 25, 1918). As Kerr states, the delay of one year "will allow time and yet not too much time to make the rather complicated arrangements that are necessary to put it [the new Faculty of Education] on any kind of a proper footing." By July 1918, Kerr reported to Tory that "I feel pretty sure now that the Faculty of Education question is quiescent" (letter, Kerr to Tory, July 15, 1918). Unfortunately, the Minister of Education, Boyle, who was enthusiastic about the prospect of the University of Alberta possessing the first faculty of

education west of Ontario, began stating publicly in early August that the new Faculty of Education would commence in the fall of 1918 (letter, Kerr to Tory, August 10, 1918).

Boyle's action forced the university to act. Kerr was empowered by Tory to offer the position of Dean of Education to McNally (telegram, Tory to Kerr, August 12, 1918). Although Tory considered this development to be the best solution to a potentially awkward situation, he continued to express apprehension over McNally's lack of a doctorate (letter, Tory to Kerr, August 13, 1918). In reply, Kerr states, "I do not think you need have any fears about McNally's success or his ability to measure up to the standard of our other appointments" (letter, Kerr to Tory, August 19, 1918).

Following Minister Boyle's indiscreet and premature announcement, the Chairman of the Board of Governors of the University of Alberta, who was also Chief Justice of Alberta, met with the Minister, along with Acting President Kerr and the Chancellor of the university. Kerr (letter to Tory, August 19, 1918) reports that Boyle was persuaded to delay the creation of the new Faculty until the next meeting of the University Senate in September. Before this meeting occurred, however, a cabinet shuffle took place, with the result that Boyle was replaced as Minister of Education by George P. Smith (letter, Kerr to Tory, September 21, 1918). Smith, who represented the riding containing the Camrose Normal School, was of the opinion that the whole matter of a faculty of education had developed too quickly. He also told Kerr, after another meeting with the Ministers of Education of the other western provinces, that the need for an additional teacher training institution in Alberta was no longer as acute as it had been (letter, Kerr to Tory, September 21, 1918). In consequence of these developments, Smith ordered immediately that nothing further be done to establish the Faculty until 1920 (letter, Kerr to Tory, September 21, 1918). Although both Tory and Kerr welcomed this development initially, Kerr reports subsequently that he feared that the opportunity for the university to gain a faculty of education might be lost, because he believed Smith, "is afraid, I suspect, of losing control of the higher teacher training" (letter, Kerr to Tory, December 23, 1918). Responding to Kerr's observation, Tory added, "there is always the danger—particularly when dealing with politicians—of the matter getting into a wrong groove" (letter, Tory to Kerr, January 16, 1919). Tory indicated that he did not believe the situation to be as precarious as Kerr feared.

Although Smith stated that he continued to support the selection of McNally as the new Dean, subsequent actions suggest that Smith, and possibly other individuals in the Department of Education, did not wish the appointment to occur. McNally's duties in the Department of Education were changed almost immediately following Smith's appointment, so that McNally found himself fully responsible for contending with the problems ensuing from the provincial government's unpopular attempt at consolidating some of the small school districts (letter, Kerr to Tory, September 21, 1918; Coutts & Walker, 1964). According to Kerr (letter to Tory, September 21, 1918) Minister Smith also reported that McNally's pursuit of a doctoral degree and his desire to visit other universities to observe how their faculties of education were structured and run would be held in abeyance until the matter of the consolidated school

districts was resolved. McNally (Coutts & Walker, 1964) states that in 1919, he was given the additional responsibility of a total review and revision of the general curriculum. McNally was also having difficulty with his dissertation.

Besides the constraints placed on his time by Minister Smith, McNally's choice of topics for the dissertation displeased Thorndike. McNally (Coutts & Walker, 1964) states, "I suggested for the dissertation a number of topics having to do with Canadian education ... As none of these involved the gathering of statistics, none was found acceptable" (pp. 42-43). McNally adds that by 1920, "I could find no reason for spending months on investigations involving mathematical data, computers,[9] and the like in which I had no interest anyway, so I regretfully cancelled my registration for the PhD degree" (p. 43). By not finishing his PhD, McNally ensured that he could not be appointed Dean of Education at the University of Alberta.[10]

Although it was noted previously that Kerr was informed that no new teacher training institution for Alberta was being considered until 1920 (letter, Kerr to Tory, September 21, 1918), the provincial government was actually proceeding with plans to erect another normal school to be located in Edmonton (Alberta Department of Education, 1921). Although the Act enabling the creation of the Normal School indicated that it could be affiliated with the University of Alberta, it also stipulated that it was to be the Minister of Education, not the university, that would control the institution (Alberta Department of Education, 1921). This action indicates clearly that the provincial government was unwilling to relinquish any control over teacher training. By the time the new Normal School opened in January 1920, it was clear that the University of Alberta would not see a faculty of education in the foreseeable future. The fleeting opportunity that Kerr had noted to Tory early in 1918 had indeed been lost (letter, Kerr to Tory, May 15, 1918).

Regrouping and New Strategies

Third Attempt: Tory's Initiative

The duplicity shown by Minister Smith prompted both Tory and Kerr to be suspicious of future government initiatives. Shortly after the announcement of the Normal School for Edmonton, Tory decided to conduct his own research to prove to the government that university-based teacher training, not state-run normal schools was the current trend in the United States. In April 1920 Tory sent questionnaires to 28 universities in the United States that offered some variety of education program.[11] Among those solicited were publicly funded institutions including: University of Illinois, Ohio State University, and University of Minnesota; privately-funded institutions included the University of Chicago and Harvard. All but two of the universities solicited replied to the questionnaire, which asked five questions specific to the configuration and administration of education programs offered by those universities. What Tory was most anxious to ascertain was whether the universities' education programs were administered primarily by the university or by the state government, and who was responsible for appointing staff. Most of the universities indicated that they alone were responsible for their education programs, including staff appointments, although the state retained responsibility for certification of teachers. This trend was congruent with Tory's image of what a faculty of education would be like at the University of Alberta.

Before the provincial government could react to Tory's study, an election was held in 1921, resulting in the defeat of the incumbent Liberal government. Although Tory and Kerr were pleased that Smith lost his seat (letter Kerr to Tory, July 20, 1921), the new government of the United Farmers of Alberta (UFA) was viewed with trepidation by Tory. Kerr presented a more positive view, stating, "I know it is hard having, so to speak, to start all over again to make converts of a lot of new men, but perhaps conversion this time will be easier" (letter Kerr to Tory, July 29, 1921). The new Minister of Education was Perren Baker who had a PhD in sociology, but preferred a career as a rancher and farmer in southern Alberta (letter, Kerr to Tory, July 29, 1921).

In spite of Tory's findings, the UFA government took no action toward the establishment of a faculty of education. A severe economic recession combined with a predilection to cater to rural constituents were the two main reasons for this (letter, Kerr to Wallace, July 18, 1928).[12] The prolonged recession prompted the provincial government to close one of the normal schools in 1923. The school in Edmonton was chosen (Coutts & Walker, 1964). This action and the government's stalling on the matter of a faculty of education agitated a number of school teachers and university professors who believed that teacher training in normal schools was inferior to what could be offered by the university (Chalmers, 1978). Beginning in 1923, and continuing through 1927, resolutions were passed at each Annual General Meeting of the Alberta Teachers' Alliance (now the Alberta Teachers' Association) encouraging both the provincial government and the University of Alberta to establish either a college of education affiliated with the university, or a full-fledged faculty of education.[13] President Tory was also taking action on his own.

Until the early 1920s the education courses offered by the university since 1912 were taught predominantly by lecturers hired by the year. During 1923 Tory, in consultation with Dr. J. M. MacEachran, Head of Philosophy and Psychology, sought the appointment of a professor who would be responsible for teaching the existing education courses and developing new ones. Tory was faced with the same problem that beset the university in 1918, namely, finding a qualified individual who was also experienced with the structure of public education in Alberta. While McNally remained in the Alberta Department of Education, he was associated with the department's stalling, and he had done no further work on a doctoral degree.[14] One school inspector, Milton Ezra LaZerte, was far removed from the administration of the Department of Education, possessed intimate knowledge of Alberta's education system, had taught recently as a lecturer at the University of Alberta, and had been pursuing a doctorate at the University of Chicago since 1920.

LaZerte (1885-1975) who worked closely with MacEachran, was hired by Tory in 1924 as a full-time lecturer in the Department of Philosophy and Psychology on the basis of MacEachran's assessment and assurance that LaZerte would complete his doctorate (Myrehaug, 1972; Chalmers, 1978).[15] LaZerte's appointment was justified to the satisfaction of both the university and the Department of Education by Tory's stating that the education courses offered were solely for the edification and preparation of the university's own graduates, not students or graduates of the normal schools (letter, Kerr to Wallace, July 18, 1928).

In 1923 Tory also struck the Committee on Teacher Training and appointed MacEachran as its chair (Council of the Faculty of Arts and Sciences, December 5, 1927). The committee was charged with formulating plans and initiating meetings with the Department of Education whereby an agreement might be reached on the matter of creating a university-operated faculty, college, or school of education.

In view of the strong resolution in favor of a faculty of education passed at the Annual Convention of the Alberta Teachers' Alliance in 1924, Tory sent another proposal for a faculty of education to Minister Baker.[16] The proposal contained Tory's findings from his 1920 survey and a detailed outline of how the university would structure and operate a faculty of education. The Department of Education responded by conducting its own replication of Tory's study and holding a special conference in November 1924, to which only selected employees of the Department of Education were invited. Although the findings of the Department's study concurred with those of Tory, no recommendations were made to establish a faculty of education. Instead, it was recommended that the solution for providing satisfactory qualifications to prospective senior high school teachers be "the provision of a College of Education under the direct control of the Department of Education" (Alberta Department of Education, 1924, p. 8). In spite of support for a faculty of education by the Alberta Teachers' Alliance and the hiring of a professor to teach education courses primarily, the university was rebuffed yet again in its attempt to establish a faculty of education. Kerr observed later that the Department of Education was unwilling to "surrender control over any part of the Teacher Training field" (letter, Kerr to Wallace, July 18, 1928).

Fourth Attempt: Tory's Last Try

To be sure, not everyone was in favor of a faculty of education. Chalmers (1978) reports that some professors in the University of Alberta did not think highly of the university containing a faculty of education. Similarly, opinions of some individuals in the Department of Education ranged from dubious to hostile toward the idea of a faculty of education (Chalmers, 1978). In most instances, aspersions were cast generally on the idea of a faculty, but in some cases LaZerte became a focus of criticism because of his position in the university.[17] In spite of such criticism, Tory and others continued to press the provincial government for permission to create a faculty of education. Further approaches to the Department of Education by the university were invariably rejected. Kerr (letter to Wallace, July 18, 1928) states, "we were not met with refusal but simply constant delay in our efforts to bring the matter to a head." Developments during 1927 finally did enable the matter to be brought to a head.

Additional support for the position of the Alberta Teachers' Alliance (ATA) came from a new group established in 1927, the Education Society of Edmonton (Chalmers, 1978). Although the ATA consisted primarily of schoolteachers, the Education Society also included professors from the University of Alberta, including LaZerte and MacEachran, employees of the Department of Education, school principals, and other school administrators. The Education Society provided further indication that there was some popular support for the idea of university-based teacher education in Alberta. Although the university's

position was bolstered somewhat by LaZerte's completion of his doctorate in 1927, the Department of Education continued to defended its old position doggedly.

At some point in 1926, the Department of Education informed Tory that its research found that university-controlled Schools or Faculties of Education in the United States were proving to be unsuccessful, with the result that many states were encouraging the creation of colleges of education separate from the universities and entirely under the control of the state (letter, Tory to President Coffman, University of Minnesota, November 1, 1927). In late 1927, in response to this claim, Tory, aided by MacEachran as Chairman of the Committee on Teacher Training, undertook another survey of faculties of education in American universities. The new questionnaire began by stating the Department of Education's position and asking whether this view was accurate. Eight other questions concerned with administrative matters were also asked.[18] All the surviving responses indicate that the Department of Education's statements concerning American faculties of education were false. For example, the President of the University of Minnesota stated, "I do not believe that there is any ground for assuming that schools of education in the universities of the United States have not been a success. I do not know of a single one that has been abandoned since it was established" (letter, Coffman to Tory, November 7, 1927).[19] In the report submitted to the Department of Education, the Committee on Teacher Training also presented 12 reasons why teacher education at the university would be superior to that offered by the normal schools. Some of the reasons listed include: highly qualified instructors (with the implication that the quality of instructors at normal schools was inferior); full university credit for courses taken; cooperation of university professors in academic training; and the use of the university library, laboratories, and residences.

Showing a spirit of compromise, the university indicated that it was willing to forgo the immediate establishment of a faculty of education. Instead, a school of education would be established under the auspices of the Faculty of Arts and Sciences. The school would be developed later into a faculty (University of Alberta, 1927). A school of education not responsible for all aspects of teacher training would not require the Department of Education to relinquish all its control in this area. Although the Department of Education seemed amenable to this idea, no action was taken.

To make matters worse, the Department of Education announced the reopening of the Edmonton Normal School (Council of the Faculty of Arts and Sciences, December 5, 1927). Although the Department of Education claimed that they were re-opening the school in response to a lengthening of the normal school session (Alberta Department of Education, 1929) the move was considered unfavorably by the university. In response to this apparent slight, the university decided to force the issue so that the Department of Education could no longer obfuscate. The *Minutes* of the Council of the Faculty of Arts and Sciences (December 5, 1927) state, "it has seemed necessary that the University should take steps to protect its own right to establish and maintain a school of education and to make clear, also, its desire to do for the teaching profession what it has done for the other learned professions." It is also reported that MacEachran received support to present the proposal for a

school of education to the Senate and to the Board of Governors of the University of Alberta (Council of the Faculty of Arts and Sciences, December 5, 1927).

The proposal for establishing a school of education was approved by the Senate on February 2, 1928 (Minutes from the meeting of the Senate, February 2, 1928). In view of the support both from within the university and from the teaching profession in the province, and in view of the likelihood that the proposal would be carried by the Board of Governors, the Minister of Education, Perren Baker, gave oral approval for a school of education to the Committee for Teacher Training by May 10 (Council of the Faculty of Arts and Sciences, May 10, 1928).[20] The Board of Governors approved the creation of a school of education, and also approved the appointment of LaZerte as its director (Board of Governors, May 15, 1928).

The School of Education was intended to provide one additional year of instruction to students graduating from the University of Alberta with either a BA or a BSc degree and who wished to teach in an Alberta high school (University of Alberta, 1928). To allow the university to complete planning and preparation, the School of Education was not launched until the fall of 1929. Some control over the preparation of teachers had finally been wrested from the provincial government through the combined efforts of some university staff, the administration of the university, the Senate, the Board of Governors, the Alberta Teachers' Alliance, and the Education Society of Edmonton .

Later Developments

School to College

No extensive change in the administrative status of the School of Education occurred until after the provincial election of 1935, which saw the demise of the UFA government at the hands of the Social Credit Party.[21] The new Premier, William Aberhart, had been a teacher and a high school principal for many years, and also became Minister of Education. Aberhart soon appointed McNally as Deputy Minister of Education. McNally claims that he advised Aberhart to sanction a plan "to bring all teacher education under the jurisdiction of the university and so increase the prestige of the teaching profession" (Coutts & Walker, 1964, p. 71).[22] Although McNally claims that Aberhart was amenable to the idea, he did not act upon it because there was no money to undertake such a change at that time (Coutts & Walker, 1964).

Continued agitation by the ATA, a more sympathetic provincial government, and an improvement in Alberta's economy enabled the university to elevate the status of the School of Education to a semiautonomous College of Education in 1939 (Kerr, 1940). At the same time, LaZerte was appointed Principal (Kerr, 1940). The change in status enabled the new College of Education to receive separate funding from the general revenue of the University of Alberta and to offer particular degrees at the master's level (Council of the College of Education, March 19, 1940; LaZerte, 1941). The College of Education was to have a short life, however.

Faculty of Education: Tory's Goal Realized

It was a tradition in Alberta for the university to bestow an honorary Doctor of Laws degree on the Premier at some point. The awarding of honorary degrees is a function of the Senate which was at that time the highest academic body of

the university. In 1941 the Senate refused to award an honorary degree to Premier Aberhart.[24] This prompted the resignation of President Kerr, and also induced some members of the provincial government to demand a complete review of all aspects of the University of Alberta (Coutts & Walker, 1964; Macdonald, 1958). The committee, referred to as the Survey Committee, was given sweeping powers. It consisted of six appointees including President Newton, Deputy Minister of Education McNally, and J.W. Barnett, the Secretary of the ATA (Macdonald, 1958). One of the first recommendations of the committee was that the College of Education should be transformed into a faculty of education as soon as possible. This occurred in 1942, 30 years after President Tory first attempted its creation.[25]

Two main factors prompted the Survey Committee to recommend the creation of the Faculty. First, the demands of the World War II resulted in the displacement of the Edmonton Normal School to smaller quarters, the Normal School building being used as a training center by the Royal Canadian Air Force (Coutts & Walker, 1964). Second, this reduction in capacity for teacher training occurred simultaneously with a severe teacher shortage caused by men enlisting in the military. Because the Camrose Normal School had already been closed in 1938, a greater responsibility for teacher training in northern Alberta fell to the College of Education, which was not designed to provide an entire undergraduate program. The creation of the Faculty of Education signified a further loss of control by the provincial government over teacher training.

In 1945, acting on one of the findings of the Alberta Post-War Reconstruction Committee, the provincial government moved to close both the Calgary and Edmonton Normal Schools and to permit the University of Alberta to be solely responsible for teacher preparation in Alberta (Chalmers, 1978).[26] Thus Alberta became the first province in Canada where all teacher education was provided by a faculty of education (Coutts & Walker, 1964; Chalmers, 1978). The certification of teachers remained under the control of the Department of Education, however.

Conclusions and Implications

From the foregoing account, we see how the factors of politics, personality, economics, qualifications, fear of, and reluctance to change all combined to delay the establishment of a faculty of education at the University of Alberta for 30 years. It is also shown that progress toward the creation of a faculty of education was reached only when the university itself controlled some of those factors and its staff and administration acted in concert with the teaching profession. Although attempts were made initially to cater to particular politicians, it was soon learned that the strategy did not work for two primary reasons. First, the tenure of politicians is frequently short-lived, and successors may not share the views of their predecessors. Second, politicians, and civil servants to a lesser degree, are primarily concerned with their own survival and the conservation of their power and control, rather than with the goals and aspirations of such institutions as universities.[27] Although it has been shown that power rests ultimately with the government in power, through its control of finance, it is up to the institutions to indicate their goals clearly and to stress their belief that the leadership and guidance for programs of study rests with

the institution, not with individual politicians or the government of the day. Above all, the tone that President Tory set, that professional disciplines such as agriculture and education were legitimate areas of study at the university, provided a clear indication of the intentions of the university, not only to the provincial government, but also to the staff of the university, professional educators in the field, and to other universities.[28] It is apparent that at the time that Tory expressed the desire to create a faculty of education he was not trying to follow a trend established at other universities, as it has been shown that education faculties were rare in North America at that time. The University of Alberta strove to be a leader, an initiator, not a follower or imitator. Although the provincial government controls the university ultimately through budget allocation, it has been shown how the university did, on occasion, through exerting its leadership against the will of the provincial government, indicate to those in power that the knowledge and competence for determining the configuration of the university rested with that institution, for example, Tory's challenging of claims by the government that university-based faculties of education were unsuccessful, and the university's proceeding with the creation of a school of education before the government sanctioned the action.

Now the Faculty of Education of the University of Alberta is faced with budget cuts and criticism from particular politicians and factions in both Alberta Education (the Department of Education renamed) and the Department of Advanced Education. The University of Alberta is once again faced with the choice of either exhibiting leadership as a counter to the will of the provincial government, or submitting to the desires of those politicians in power, thus conceding that leadership rests with the provincial government.

Although it may be tempting for the administration of the Faculties of Education and the universities experiencing such financial constraints to diminish and to close particular programs in the interest of budgetary constraint in the belief that such action will be perceived as appropriate cooperation by the provincial government and will be rewarded, such an approach is both divisive and demonstrative of an ignorance of more successful strategies used by administrators in the past.[29]

To ignore the experiences of the past, or to dismiss them as irrelevant, is also foolhardy because many of the factors that impinged on President Tory and the University of Alberta in its early years continue to exist and to affect the University of Alberta and probably other state-funded universities in North America, although particular circumstances differ from place to place, and some particular aspects are different from what they were 80 years ago (Birnbaum, 1992; Cameron, 1986). If we choose to consider the struggles that faced our predecessors in their attempts to create the universities that we have today only as interesting but quaint stories of a different era, and if we maintain that universities are creatures of governments and that we must be beholden to them, then instead of building on what our predecessors left us, we risk losing much of what they achieved. In such circumstances, instead of devoting our efforts to improving, expanding, and enhancing our legacy and trust, we risk spending the future struggling to recreate and recapture what we allowed to be taken away. The work of Tory and others indicates that the solution is neither confrontation nor the administration of the university impressing its

own *politically correct* will on the rest of the university in the guise of *team spirit* (Lively, 1992). As shown in the past, a better approach is a concerted effort by the university administration to work together *with* its staff, professional organizations, and the relevant branches of the provincial civil service, to develop and to present clear goals to the politicians and the sovereign people and to provide strong leadership.[30] In this way the experiences of the 30 years leading to the creation of the Faculty of Education will not be lost.

Notes

1. The author acknowledges the kind and generous assistance of Dr. H.T. Coutts, the third Dean of the Faculty of Education.
2. Tory, who had come from McGill University in Montreal, was following the McGill model by considering both agriculture and education as legitimate areas of study at a university. In spite of the McGill model, the inclusion of the study of education in state-funded universities in North America was rare before the World War I. Edmonson and Webster (1925) report that although the University of Indiana began a teacher training program in 1853, there were only 18 state-funded universities in the United States offering some courses or programs in education by 1914. In Canada, similarly, few universities besides McGill in Montreal and Queen's in Kingston offered programs in education before the World War I. The progress in Canada was slower than that in the United States. The slower pace in Canada was abetted in part by the views of some individuals that the best courses in education were those offered in the United States. Indeed, as late as 1937, the Saskatchewan Department of Education's (1938) *Annual Report* states that many teachers sought upgrading, "not only in Canadian universities, but in the great educational centres in the United States" (p. 35)
3. The qualifications of individuals employed in normal schools in other provinces appears to have been similar (Gregor & Wilson, 1984; see also Saskatchewan Department of Education, 1906-1940; Ontario Ministry of Education, 1905-1939)
4. The Senate and the Board of Governors did not consider the creation of a department because of cost. Instead, it was approved "that the subject of Education be added to the course of study in the third and fourth years [for students enrolled in the Faculty of Arts and Sciences]" (Board of Governors Executive Committee, Minutes, April 25, 1912).
5. This work represents McNally's memoirs as transcribed by Dr. Herbert T. Coutts and Dr. Bernal E. Walker, the former being the third Dean of the Faculty of Education, the latter being a long-time professor in the Faculty.
6. The correspondence cited is located in the University of Alberta Archives in two main collections: letters and telegrams between Kerr and Tory are located in the Henry Marshall Tory Papers; other letters are located in the Robert C. Wallace Papers.
7. Tory's duties in this capacity required him to remain in England until 1919.
8. The potential problem of contending with Miller was avoided by his departure in May 1918 to the University of Minnesota (letter, Kerr to Tory, May 15, 1918). This development caused Tory to write, "When I saw that he had gone to Minnesota, I rejoiced" (letter, Tory to Kerr, August 13, 1918).
9. When using the term *computers*, McNally is referring to individuals who compile data and perform particular calculations with it.
10. Although McNally eventually became Deputy Minister of Education in 1935, the only doctorate he obtained was an honorary Doctor of Laws bestowed by the University of Alberta in 1946 for his long service to education in the province.
11. The questionnaire and most of the responses are located in the Henry Marshall Tory Papers, University of Alberta Archives.
12. The deleterious effects of the recession and political patronage on existing educational institutions and the development of new programs were also experienced in other provinces (Gregor & Wilson, 1984).
13. Copies of these resolutions are located in the Robert C. Wallace Papers, University of Alberta Archives.
14. Although in his memoirs McNally (Coutts & Walker, 1964) does not mention anything about his ill-fated near appointment as Dean of the Faculty of Education in 1918, he gives the impression that he was always supportive of the idea of a faculty of education. This position is disputed by LaZerte (Myrehaug, 1972).

15. LaZerte completed his doctorate in 1927 under the direction of Charles H. Judd at the University of Chicago.
16. Portions of this proposal plus the report of a consequent conference conducted by the Department of Education are located in the Robert C. Wallace Papers, University of Alberta Archives.
17. This was especially true in the case of Charles Sansom, who worked in the normal schools, and who held the opinion that universities were institutions concerned with esoteric and largely useless research rather than with practical applications of knowledge and skill (Sansom, 1927a; 1927b). Rather than using a reasoned approach, Sansom relied on invective not only in his opposition to a faculty of education, but also in his criticism of LaZerte's promotion of research in education (Sansom, 1928).
18. The questionnaires and most of the responses are located in the Henry Marshall Tory Papers, although some responses and summaries are located in the Robert C. Wallace Papers.
19. The matter was further exacerbated when it was learned that the Department of Education planned to construct a new normal school building immediately south of the campus of the University of Alberta. This building, now named Corbett Hall, is now part of the university.
20. The Edmonton Normal School was reopened during 1928. It was closed subsequently in 1933, primarily because of the effects of the Great Depression.
21. From correspondence in the Robert C. Wallace Papers, it seems that in early 1935 President Wallace intended to elevate the status of the School of Education to a separate Department of Education within the Faculty of Arts and Sciences. Although LaZerte acknowledged this change in correspondence, no mention of such a change appears in any official report. Given the extreme financial conditions of the time, it is likely that the change was not approved, because the change was dependent on the budget provided to the university by the provincial government (letter, Wallace to LaZerte, January 14, 1935).
22. McNally's action seems out of character with the previous actions of the Department of Education. It may be, however, that in spite of LaZerte's sentiments (Myrehaug, 1972) McNally was always in favor of this, but kept his own ideas secret when the largely unsympathetic UFA was in power.
23. Enough money could be found, however, to reopen the Edmonton Normal School in 1935. The School remained open until its final closure in 1945.
24. As a politician, Aberhart came to be hated vehemently by several sections of the public as his tenure progressed, and many individuals expressed their hatred openly in tangible ways.
25. Although Tory was President when approval was given for the School of Education, he resigned from the University of Alberta in 1928 to assume the directorship of the National Research Council.
26. Although McNally (Coutts & Walker, 1964) claims that he was the driving force behind this initiative, the consolidation of teacher training at the University of Alberta could not have taken place without the forbearance of the Social Credit government and the cooperation of LaZerte and the administration of the University of Alberta.
27. Events at the University of Manitoba show that this phenomenon is not unique to the University of Alberta. Gregor and Wilson (1984) state that the malfeasance of the Chairman of the Board of Governors of the University of Manitoba, which was discovered eventually by 1932, delayed that university's building program and the establishment of a faculty of education. Extensive changes to the administrative structure were made subsequently to help ensure that a repeat of such a situation would be unlikely. Sensing a loss of public support and a loss of trust by many individuals at the university, the Manitoba provincial government attempted to regain both support and credibility by increasing the funding of the University of Manitoba so that it was able to establish a faculty of education by 1935 (Gregor & Wilson, 1984, p. 113).
28. Barnard (1938) agrees that the leadership shown by a university or college president is a most important element in determining effectiveness and the future direction of the institution. This view is also congruent with the more recent findings of Miles and Cameron (1982) and Cameron (1986).
29. Kerr (1984) summarizing recent surveys of United States colleges and universities, stresses that both the administration and the Board of Governors should strive to keep "partisan politics out of the university, just as the university itself should stay out of partisan politics" (p. 83).

30. Considering contemporary developments and findings from recent surveys conducted largely in the United States, both Cameron (1986) and Birnbaum (1992) recommend the adoption of cooperative approaches by university and college administrators.

References

Alberta Department of Education. (1907). *Annual report of the Department of Education of the Province of Alberta 1906*. Edmonton, AB: Government Printer.

Alberta Department of Education. (1921). *Fifteenth annual report of the Department of Education of the Province of Alberta 1920*. Edmonton, AB: King's Printer.

Alberta Department of Education. (1924). *Report of special conference re professional preparation of high school teachers*. Edmonton, AB: Author.

Alberta Department of Education. (1929). *Twenty-third annual report of the Department of Education of the Province of Alberta 1928*. Edmonton, AB: King's Printer.

Barnard, C.I. (1938). *The functions of the executive*. Cambridge, MA: Harvard University Press.

Birnbaum, R. (1992). *How academic leadership works*. San Francisco, CA: Jossey-Bass.

Board of Governors, University of Alberta. (1912-1928). *Minutes*. Edmonton, AB: Author.

Cameron, K.S. (1986). A study of organizational effectiveness and its predictors. *Management Science, 32*(1) 87-112.

Chalmers, J.W. (1978). *Gladly would he teach: A biography of Milton Ezra LaZerte*. Edmonton, AB: ATA Educational Trust.

Chalmers, J.W. (1992, Winter). Development of Alberta's first faculty of education. *Alberta History*, 12-16.

Council of the College of Education, University of Alberta. (1940). *Minutes*. Edmonton, AB: Author.

Council of the Faculty of Arts and Sciences, University of Alberta. (1927-1928). *Minutes*. Edmonton, AB: Author.

Coutts, H.T. (1979). *Yesterday tomorrow's dream and today's vision*. Edmonton, AB: Unpublished speech.

Coutts, H.T., & Walker, B.E. (1964). *G. Fred: The story of G. Fred McNally*. Don Mills, ON: Dent (Canada).

Edmonson, J.B., & Webster, A.H. (1925). *Policies and curricula of schools of education in state universities*. Washington, DC: Department of the Interior, Bureau of Education.

Gregor, A., & Wilson, K. (1984). *The development of education in Manitoba*. Dubuque, IA: Kendall/Hunt.

Kerr, C. (1984). *Presidents make a difference: Strengthening leadership in colleges and universities: A report of the Commission on strengthening presidential leadership*. Washington, DC: Association of Governing Boards of Universities and Colleges.

Kerr, W.A.R. (1940). Report of the president. *Report of the Board of Governors of the University of Alberta for the year ending June 30, 1940*. Edmonton, AB: University of Alberta.

LaZerte, M.E. (1941, Summer). College of Education administers B. Educ. Degree. *College of Education Alumni Society Quarterly Bulletin*, p. 2.

Lively, K. (1992, November 4). Deep cuts in state spending produce a bitter mood at San Diego State. *The Chronicle of Higher Education*, pp. A21-A23.

Macdonald, J. (1958). *The history of the University of Alberta 1908-1958*. Edmonton, AB: The University of Alberta.

Miles, R.H., & Cameron, K.S. (1982). *Coffin nails and corporate strategies*. Englewood Cliffs, NJ: Prentice-Hall.

Myrehaug, D.M. (1972). *M.E. LaZerte: Contributions to teacher education in Alberta*. Unpublished master's thesis, University of Alberta.

Ontario Ministry of Education. (1906-1939). *Annual reports*. Toronto, ON: King's Printer.

Patterson, R.S. (1979). History of teacher education in Alberta. In D.C. Jones, N.M. Sheehan, & R.M. Stamp (Eds.), *Shaping the schools of the Canadian west*. Calgary, AB: Detselig.

Sansom, C. (1927a). The training of teachers I. *ATA Magazine, 8*(4), pp. 17-18.

Sansom, C. (1927b). The training of teachers II. *ATA Magazine, 8*(5), p. 7.

Sansom, C. (1928). More about research. *ATA Magazine, 8*(6) pp. 15-16.

Saskatchewan Department of Education. (1906-1940). *Annual reports of the Department of Education of the Province of Saskatchewan*. Regina, SK: King's Printer.

Saskatchewan Department of Education. (1938). *Annual report of the Department of Education of the Province of Saskatchewan, 1937*. Regina, SK: King's Printer.

Tory, H.M. (1912). Report of the President. *Report of the Board of Governors of the University of Alberta for the year ending June 30, 1912*. Edmonton, AB: University of Alberta.
University of Alberta. (1912). *Annual calendar of the University of Alberta 1912-1913*. Edmonton, AB: Author.
University of Alberta. (1927). *Report of the committee on teacher training*. Edmonton, AB: Author.
University of Alberta. (1928, September 24). *School of Education established at the University*. Press release.
University of Alberta. (1929). *Report of the Board of Governors of the University of Alberta for the year ending June 30, 1929*. Edmonton, AB: Author.

The Alberta Journal of Educational Research *Vol. XXXIX, No. 3, September 1993, 335-346*

Ken Pisichko
Seven Oaks School Division, Winnipeg

Manitoba's GSE: An Attempt at Resolving the Fiscal Equity Dilemma

The Manitoba Government Support to Education (GSE) Program of public school funding was implemented in 1985 to enhance equity in Manitoba schools. This article provides the first statistical evaluation of Manitoba public school finance equity provisions. Selected fiscal disparity and wealth neutrality measures tested for the equitable distribution of revenues appropriated to school divisions by the province. Each of the equity measures was applied to the total aggregate per-pupil revenue provided by the previous Education Support Program (ESP) and the (GSE). The results of the ESP analysis formed a baseline for evaluating changes in fiscal equity during the first three years of the GSE. In addition, the study developed models that indicated (through regression analysis) a relationship between per-pupil revenues allocated by the province and the enrollment size of a school division and its location in the province.

Afin de promouvoir l'équité financière pour les écoles au Manitoba, le Manitoba Government Support to Education (GSE) Program a été institué en 1985. Cet article présente la première évaluation statistique des provisions d'équité pour le financement des écoles publiques au Manitoba. Certains moyens de reconnaître la disparité fiscale et la neutralité de la répartition des richesses ont permis de mesurer la distribution équitable des revenus appropriés pour les différentes divisions scolaires à travers la province. Chacun des moyens d'équité a été examiné en fonction des revenus totaux par-élève permis par les programmes Education Support Program (ESP) et (GSE). Les résultats de l'analyse du ESP ont formé une ligne de base afin d'évaluer les changements dans l'équité financière pendant les premières trois années du GSE. En plus, cette recherche a développé des modèles qui ont indiqué (par l'entremise de l'analyse de regréssion statistique) qu'il y a une relation entre les revenus par-élève alloués par la province et le taux d'inscription d'élèves d'une division scolaire ainsi que son emplacement géographique dans la province.

School finance reform has been a controversial topic in the United States for several decades. Such controversy has manifested itself in court cases at the state and national level. Typically, litigation tested the constitutionality of state school finance provisions. These court cases have resulted in a variety of new public school finance provisions and often in additional litigation to test the constitutionality of the subsequent school finance provisions. On the other hand, Canada has not experienced such challenges to school finance provisions. Lawton (1979) suggested a socialistic political philosophy as enunciated by Hartz was the major factor driving this dearth of litigation. Lake (1983) suggested that under Canadian law there was no provision for the testing of the constitutionality of school finance equity provisions. MacKay (1986) provided further legal opinion when the Canadian constitution (1982) was scrutinized. In addition to these citations I have noted the absence of

---

Ken Pisichko is a classroom teacher whose research interests include funding, collective bargaining, and policy analysis.

Canadian school finance litigation in the decade since the Canadian Constitution or Charter was proclaimed. Nevertheless, there have been various reforms to the school finance provisions in the provinces and territories of Canada. Although Bezeau (1979) and Lawton (1979) have discussed various fiscal equity indicators in their studies of Canadian school finance provisions, I have noted that such quantitative indicator usage is limited in Canada and is seldom quoted in Canadian studies. This analysis utilized quantitative indicators with a view to evaluating the fiscal equity provisions of the Government Support to Education Program.

Manitoba School Finance Reforms

Most public schools in Manitoba were of the one-room variety until the late 1950s when several royal commissions and studies recommended substantial changes to the organizational structure and funding of Manitoba's school districts. The MacFarlane Report of 1959 (Manitoba Government, 1959) resulted in the first major changes to school funding by including substantial financial incentives for establishing school divisions with high schools. This first move toward consolidation was followed by the Michener Commission in 1964 (Manitoba Government, 1964), which recommended the establishment of unitary school divisions and further financial incentives. The move toward school district reorganization was completed by the early 1970s. School divisions in Manitoba adopted a unitary organizational structure, with an elected board of trustees and an appointed superintendent acting as Chief Executive Officer. The exception to this pattern is Frontier School Division in northern Manitoba, which has no elected board of trustees, a small property tax base, and an Official Trustee and Superintendent appointed by the Minister of Education. Several school districts remain as anomalies in the province, however, on military installations, in remote areas, or in communities with a limited tax base.

Following the Michener Report, there was a lapse of several years between school jurisdiction reorganization and additional school finance changes. Another study was initiated by, and presented to, the Manitoba government in 1967, and resulted in a funding provision called the Foundation Program. This document provided for the first time in Manitoba a written and legislated commitment to the basic equity principles in public school funding.

By 1980 there was a great deal of concern regarding the Foundation Program. The provisions had remained essentially unchanged since its implementation in 1967. In spite of annual increments to the program funding components, the spread between provincial support and school board expenditures widened as the program became less capable of responding to new situations and needs. In response, the Manitoba Department of Education introduced the Education Support Program (ESP) in 1981 as a funding revision. This program retained the basic principles of providing for most school board expenditures from the provincial treasury while preserving local autonomy. The ESP also provided for increased equalization over the entire school system, for increased school funding from general treasury revenues and away from real property taxation, and for further expansion and development of programs and services. Legislation also provided for future require-

ments and plans and for revenue loss due to declining enrollment (Manitoba Government, 1981).

The ESP provided more equalization by increased support from the provincial treasury. Furthermore, local school boards could theoretically reduce local property taxation through a reduced special levy. As Jefferson indicated (1985), this projection did not materialize because some school boards maintained their special levy mill rate while using their traditional ESP funding for other educational expenditures. The ESP also provided a framework for evaluation and review, creating for the first time in the province a systematic review of Manitoba's funding provisions.

In 1981 a change in government with a different political philosophy led to an Education Finance Review Committee. The committee followed the same procedures utilized by previous royal commissions, holding public meetings and receiving briefs and written submissions throughout the province. The committee report entitled *Enhancing Equity in Manitoba Schools* was presented to the Minister of Education in 1983, with enactment of the Government Support to Education Program (GSE) in the following year. The Minister of Education provided the following description of the new provisions:

> In order to ensure that divisions will not receive less block and equalization support in 1985 than they received in 1984, some divisions receive funding by the variable block method.
>
> Block and equalization support is the total of the support provided to the division in the education support program and other support for the following five components:
>
> basic operating support,
> pupil support,
> extra operating support,
> eligible expenditure support, and
> equalization supplement
>
> or, if greater, the total of the five components divided by the 1984 eligible enrollment with the result multiplied by the 1985 eligible enrollment. (Manitoba Government, 1985, p. 2)

Local autonomy and the special levy were retained as was the grandfathering option provided by the Variable Block Support. Unknown to the government, the retention of this option would result in some circumvention of the expressed policy aimed at enhancing equity. The grandfathering provision was provided in the revised Guarantee Support, which provided a cushion against revenue loss. Funding revisions also increased eligible enrollments by weighting that were not present in the ESP provisions. Under Block Support and Equalization Supplements, the weighted enrollment provision increased eligible enrollments by weighted factors related to declining enrollment, economies of scale, northern allowance, and teachers' qualifications and experience.

It was evident that the funding provisions were being fine tuned on a more frequent basis in an effort to reduce funding inequities. This process also became part of the mandate of the Advisory Committee on Education Finance established by the Minister of Finance in 1988.[1]

The result of Manitoba school finance reform through the GSE was a plan that subscribed to the concepts of horizontal and vertical equity and to equal educational opportunity, while still allowing for local autonomy and control. School finance literature is rich with equity measures; however, there has been no application of these measures to Manitoba school finance. Furthermore, Manitoba has published no documents regarding equity changes from a quantitative point of view. The Minister of Education has only released qualitative and political statements regarding the equality of the GSE provisions. As a result, this study utilized several measures common in school finance literature to evaluate the extent to which fiscal equity was achieved by the Manitoba Government Support to Education Program.

Research Questions

The study was driven by the following research questions:

1. What is the difference in fiscal disparity of provincial revenues of Manitoba school divisions between the base year, 1984, and the years 1985 through 1987 when the GSE program was in operation?
2. What is the difference in the wealth neutrality of provincial revenues of Manitoba school divisions between the base year, 1984, and the years 1985 through 1987 when the GSE program was in operation?
3. What is the relationship between school district size and the equitable distribution of provincial revenues for Manitoba school divisions for the years 1985 through 1987 when the GSE program was in operation?
4. What is the relationship between school division geographic location and the provincial revenues for Manitoba school divisions for the years 1985 through 1987 when the GSE program was in operation?

Study Design

This study was structured by the following sequence of steps.

Population and Sampling

Data were collected on all 47 school divisions in Manitoba. Only school divisions have the use of local property taxation as a significant support for the local school board. As this study investigated the impact of the province's funding on the local boards, the locally generated revenues were not used. Furthermore, the half-dozen operating school districts were not included for the following reasons. Manitoba school districts have a virtually nonexistent residential property tax base because of their locations on or adjacent to Crown lands, military installations, and in isolated regions of northern Manitoba. Schools operated on Indian Reserves, as well as several private independent schools were not included either because they did not obtain revenues from the province. Inclusion of such school districts would create difficulties in analyzing wealth neutrality. In addition, such anomalies represent a tiny proportion of the entire student population and revenue picture in Manitoba.

Data and Sources

Aggregated provincial revenue figures for all pupils were analyzed for the two funding programs studied: the Education Support program in 1984 to establish a baseline, and the Government Support to Education Program from 1985 through 1987. The following demographic data were also utilized: name and

identifier number of each school division, eligible enrollments, equalized assessments per pupil (BAPP), and school division location identifier.

All data were obtained from the Public Schools Finance Board of the Manitoba Department of Education, except the location identifiers which were obtained from the Manitoba Teachers' Society.

Data Treatment

Selected fiscal disparity and wealth neutrality measures tested for the equitable distribution of revenues appropriated to Manitoba school divisions by the province. Each equity measure was applied to the total aggregate per-pupil revenue as provided by the Education Support Program (ESP) and the Government Support to Education Program (GSE).

Fiscal disparity was measured by variance, the coefficient of variation, and by permissible variance. These three measures were calculated using the formulae discussed by Berne and Stiefel (1984). This approach paralleled studies conducted in various parts of Canada and the United States by Carruth (1980), Jacobsen (1986), Jefferson (1982), and Schmidt (1988). These three measures are now discussed briefly.

*Variance.*This measure is the average of all squared differences from the mean. It was calculated by summing the squared deviation of each per-pupil revenue value in the Manitoba school division population and then dividing by the total population size. Variance describes the dispersion of sample values about the sample mean value. Reduced variance is desirable in this study because it would indicate a reduction of fiscal disparity through reduced variation in provincial revenue distribution to Manitoba school divisions.

Coefficient of variation. This measure of revenue dispersion was calculated by dividing the standard deviation (the square root of the variance) by the population mean. The decimal fraction was then multiplied by 100 to give a percent. In this study this measure expressed the standard deviation per revenue dollar as a percent of the general mean revenue dollar value of the population under investigation. The investigation utilized the pupil unit for analysis as recommended by Berne and Stiefel (1984). As with variance, the smaller the coefficient of variation, the smaller the revenue dispersion. As Schmidt indicated, an increasing coefficient of variation indicated more variation in the distribution of provincial revenues and thus more discrimination on the basis of school division wealth (BAPP) in the allocation of provincial revenues.

Permissible variance. This measure is based on the revenues below the median value and is used to indicate how low revenue school divisions are progressing toward the concept of equal educational opportunity. Permissible variance is also called the McLoone Index in memory of its developer Dr. Eugene McLoone. The McLoone Index was calculated by summing the total per-pupil revenues below the median and dividing this numerator by this total plus the additional revenues needed to bring all students concerned up to the median revenue. The index varies between 0.00 and 1.00. A value of 0.00 would represent a situation of perfect inequity of the distribution of per-pupil revenue, whereas a value of 1.00 would represent a situation of perfect equity in per-pupil revenue distribution as related to equality of educational opportunity.

Fiscal (wealth) neutrality was measured with the bivariate Gini coefficient (Hickrod, 1983). The data were sorted in ascending order of property valuation (BAPP), and the resulting computation of the Gini coefficient was based on aggregated provincial per-pupil revenues plotted against property valuation (BAPP). The data plots indicated classical Lorenz curve behavior because the data curves did not cross the diagonal. This behavior confirmed the use of the Hickrod analytical technique as appropriate for wealth neutrality analysis.

The relationships in research questions three and four were investigated through multiple and partial regression analysis. In this study, enrollment was a continuous independent variable. The eight locations of Manitoba school divisions were categorical independent variables and were designated by dummy variables. If a school division was in a particular location, its dummy variable was designated as $D_i=1$; otherwise its dummy variable was designated as $D_i=0$. In this manner the eight locations were designated with seven dummy variables D_1 through D_7. In addition, an interaction independent variable was obtained by multiplying the dummy variable for a school division by its enrollment variable. In the analysis the independent variable was per-pupil revenue. The location and enrollment variables were used because they were driven by the research questions. In discussion with school trustees and teachers in the province, I decided to test the impact of these variables on the allocated revenues. The GSE description indicated that there should be no effect on revenue allocated to a pupil regardless of the size and location of the school division. The only exception to this was the specific northern allowance allocated to the three northern school divisions.

Four models were constructed to force certain variables into a regression equation. Each model was described by a regression equation having per-pupil revenue as the dependent variable, and having various independent variables. Table 1 shows the four regression models in terms of their independent variables.

Model one is the full model and the others are reduced models. All models have a coefficient of determination (r^2) that indicates the percent of the model explained by the independent variables. This coefficient thus measures how good a particular model was. Individual models were evaluated by hypothesis testing where the null hypothesis H_0 was represented by models two, three, and four. This null hypothesis was tested against the alternative hypothesis H_1 which was always represented by model one. With each analysis an *F* statistic was calculated and compared with critical values at the .10 confidence level.

Table 1
Independent Variables Associated With Models Tested in Manitoba School Finance Equity Study

Model	*Independent Variables*
1	Location + Enrollment + Interaction
2	Location + Enrollment
3	Location
4	Enrollment

The .10 level was used rather than the more common .05 level because this study used the total population of school divisions rather than a sample of Manitoba school divisions. Such methodology selected the best model for each year of available data.

The effect of location was further analyzed through partial regression as discussed by Green (1978) and Mendenhall (1986). The result of this procedure was a partial independent variable $_i$ that was plotted to illustrate the effect of each dummy variable on the functional relationship. Table 2 indicates the associated $_i$ for each school division location.

Findings of Study

Research Question One: Effects on Fiscal Disparity.

The stated purpose of the finance reform resulting in the GSE intended to increase equity and consequently to reduce fiscal disparity. Data analysis did not generally support this projection. Table 3 shows the indicator results for the years in question.

The first year of the GSE indicated a fourfold increase in the baseline variance of $41,978 to the first-year GSE variance of $153,526. The next two years of the GSE indicated smaller variance increases of approximately $25,000 per annum. These values in the variance indicate that the first three years of the GSE program increased fiscal disparity of per-pupil revenue distribution in Manitoba.

A similar increase in fiscal disparity was indicated by the coefficient of variation. An increase of 73% occurred between the values of the baseline year and the first year of the funding reform. Continued but smaller increases continued in the next two GSE years as well.

The McLoone Index showed a variable effect on vertical equity when the GSE was implemented. The change in value from .9479 in the baseline year to .9431 and .9410 in the first two years of the GSE program indicates that vertical equity was reduced in those years when compared with the previous ESP funding provisions. However, in 1987 the McLoone Index of .9513 indicated an increase in vertical equity above the baseline value. Although vertical equity was finally enhanced by the GSE in 1987, it is unlikely that the overall effect on vertical equity was positive over the first three years. The mean value of the McLoone Index for the three years of the GSE was calculated at .9451. This is

Table 2
Partial Regression Variables for Manitoba School Division Locations

Location	*Partial Independent Variable*	
North Central	1	i_1
Interlake	2	i_2
South East	3	i_3
South Central	4	i_4
South West	5	i_5
Parkland	6	i_6
Northern	7	i_7
Unicity	8	i_8

Table 3
Fiscal Disparity Indicators 1984-1987

Year	Program	Variance	Coefficient of Variation	McLoone Index
1984	ESP	$41,978	8.14	.9479
1985	GSE	$153,526	14.06	.9431
1986	GSE	$179,298	14.71	.9410
1987	GSE	$204,580	14.90	.9513

approximately 1% less than the baseline ESP value of .9479. It appears that this leads to the interpretation that the GSE led to a small reduction in overall vertical equity at a time when more equity was anticipated, even though the final year of the study showed an increase as compared with the baseline year.

Research Question Two: Effects on Fiscal (Wealth) Neutrality.
The bivariate Gini coefficient indicated a continued and increased disproportionate per-pupil revenue distribution from the GSE funding provisions. The Gini coefficients in question are indicated in Table 3.

As indicated, the GSE finance reform caused a first-year increase of 18% in the Gini coefficient followed by increases of less than 2% in each of the next two GSE program years. Property-poorer school divisions were thus placed in a situation of continued disproportionate reductions in per-pupil revenues from the Manitoba treasury, whereas property-richer school divisions received disproportionate increases in their revenues. Although the total monies allocated to each school division by the province are based on the total student population, the Gini coefficients indicate that wealth neutrality was not achieved by the GSE data analyzed, and that because the poorer school divisions received disproportionately less funding than anticipated under the principles of wealth neutrality, the students not only were not given their due, but the local taxpayers were not given their due either.

Research Questions Three and Four: Relations of Revenue to Location and Enrollment Size
The multiple variable regression analysis utilized per-pupil revenue as the dependent variable while three independent variables were tested in the model building stage of the analysis. Eligible enrollment interacted with school division locator dummy variables to produce a third interaction variable. The four models constructed are indicated in Table 1.

Testing the full model using the null hypothesis versus the alternate hypothesis *F* test approach produced consistent results for the four years in question. In spite of significance for all four models in each year of the study, the F tests consistently indicated that Model One expressed the best relationship between the dependent and independent variables for each year in question. Each year, the $F_{calculated}$ always exceeded the $F_{critical}$ at the .10 level of significance. The 90% confidence level (or .10 level of significance) was selected to reflect the use of the total Manitoba school division population in the analysis. Because the r^2 values for Model One varied increasingly from .711 in 1984 through .856 in 1987, the strength of this model was also indicated by the

independent variables explaining from 71.1% to 85.6% of the variance in per-pupil revenue.

For each year, analysis of Model One variables indicated that the relationship between per-pupil revenue and location dummy variables D_1 through D_7 was most significant. In all years the "b" values for location dummy variables were greater than 124.141 while the "b" values for the enrollment variables were .009 or less and interaction variables were always less than 2.297. Based on these results a partial regression was conducted based on the full model coefficients. This isolated and evaluated the effect of each location on the relationship between per-pupil revenue and enrollment. The partial regression produced a linear expression for each of the locations of the form $y=mx+b$ where

y = per-pupil revenue for school division,
m = slope of linear relationship,
x = enrollment size of school division, and
b = intercept value of per-pupil revenue.

Table 4 summarizes the partial regression analysis of Model One in terms of the slope m and the per-pupil intercept b values. The table does not include x values for the enrollment because the relationship is explained with the values of m and b.

These values of slope and intercept provide several interesting aspects to the concept of equal treatment of school divisions by the GSE funding reforms. The slopes are of particular interest when grouped according to city (Winnipeg), northern, and rural regions. Six of the regions are rural, and it is noted that the partial regression slopes for these regions were usually negative for the four years of the study. This negative slope indicates less per-pupil revenue from the GSE as rural school division enrollment increases. On the other hand, the slope for the northern region was always positive and increasing in value from a low of .316 in 1984 to a high of 2.451 in the last year of the study. These slope values seem to indicate that government claims of more equal treatment to school divisions by the GSE reforms are not justified.[2] Proponents of the economy of scale would recognize the impact of the negative slope; however, the negative values applied consistently to the rural areas, only once to the city of Winnipeg, and never to the northern school divisions. Reasons for this discrepancy could not be investigated by the research methodology used at the outset of the study and came as a total surprise. The increased slope of the northern region in the four years also shows that each eligible student provided $2.13 more to the northern school division in 1987 than was provided in 1984. This observation arises from the slope increase from .316 in 1984 to 2.451 in 1987. Because the partial regression is $y = mx + b$, the mx component (related to enrollment x) has increased by $2.13. The only situation when this trend in rural location was not followed occurred in 1985 when the Parkland region had a slope of +.125. Metropolitan Winnipeg had slopes that were overall much closer to zero when compared with the other regions of the province. This indicates that student enrollment in the 10 Winnipeg school divisions had a lessened effect on per-pupil revenue as compared with the effect in other regions, and particularly when compared with the northern regions. Causes for these interrelationships between per-pupil revenue and enrollment for

Table 4
Partial Regression Analysis of Model One 1984-1987

Year	*Region*	*Slope = m*	*Intercept = b*
1984	North Central	(.043)	$2587.05
	Interlake	(.041)	$2599.09
	South East	(.136)	$2894.37
	South Central	(.433)	$3186.09
	South West	(.075)	$2702.58
	Parkland	(.074)	$2616.68
	Northern	.316	$2029.38
	Winnipeg	.009	$2402.09
1985	North Central	(.060)	$2867.69
	Interlake	.125	$2642.40
	South East	(.184)	$3296.44
	South Central	(.458)	$3429.67
	South West	(.073)	2855.20
	Parkland	(.252)	$3231.55
	Northern	1.636	($1.92)
	Winnipeg	.002	$2715.36
1986	North Central	(.210)	$3038.52
	Interlake	(.236)	$3218.11
	South East	(.325)	$3484.78
	South Central	(.574)	$3525.96
	South West	(.201)	$2981.20
	Parkland	(.567)	$3733.78
	Northern	2.179	($1079.20)
	Winnipeg	(.118)	$2836.40
1987	North Central	(.084)	$3187.83
	Interlake	(.152)	$3494.21
	South East	(.235)	$3722.40
	South Central	(.437)	$3633.12
	South West	(.092)	$3138.50
	Parkland	(.501)	$3987.66
	Northern	2.451	($1206.09)
	Winnipeg	.000	$3014.36

each location would be conjecture as this stage. Regression analysis in this study could not investigate causes.

Conclusions

The fiscal equity provisions of the Manitoba school finance reform effected in 1985 were expected to enhance or maintain equity. Increased funding and grandfathering were implemented to promote this situation. Quantitative analysis of province-provided per-pupil revenues from the GSE program resulted in criticism with regard to several aspects of fiscal equity. It became evident that providing more money did not facilitate equity enhancement because horizontal equity indicators showed equity reduction with implemen-

tation of the GSE. In addition, values of the McLoone Index indicated that funding revisions did little to improve the overall "unequal treatment of unequals" as the average of the index was less than the baseline value in 1984. It is a moot point that the final value in 1987 was higher than the baseline value, as McLoone Index values fell for the first two years of funding revision before they eventually rose above the baseline value.

Further weakness occurred with regard to property wealth and provincially allocated per-pupil revenues. Increasing Gini coefficients indicated that fiscal (wealth) neutrality was reduced by the implementation of the GSE and resulted in below median or property-poor school divisions receiving a disproportionate revenue share that continued to be reduced every year. Property assessment has resulted in considerable controversy in Manitoba. It was not possible to evaluate the consistency of property evaluation in the province. Furthermore, since this study was conducted, a new approach to Manitoba real property assessment has been implemented and has resulted in much lobbying and nearly a taxpayer revolt. The situation is far from settled. Because the same assessment data were used by the Public Schools Finance Board, the validity of the use of these data was not questioned for this study.

Except for the final year of this study, neither the ESP nor the GSE had regional funding provisions or enrollment provisions. Additional funding provisions such as Small Schools Grants were provided when idiosyncratic needs arose. In spite of these grants the partial regression slopes of rural schools continued to decline, and multiple variable analysis continued to indicate a definite relationship between the school division location and the provincial per-pupil revenue the division received. Apparently the per-pupil weightings utilized by the province did not adequately overcome per-pupil revenue differences between regions. The grandfather clause may have been a partial cause of such inequity. Unfortunately, this study did not evaluate the influence of the grandfather clause.

Factors such as transportation cost and commodity supply and demand variation will, by their nature, cause variation in the cost of an equivalent education throughout the province. The funding provisions do not clearly address such cost variation. It would be enlightening to analyze the influence of such costs on the revenue distributed by provincial funding programs. The difficulty would be in identifying the exact costs and in obtaining data that was consistently reliable and valid. In this study only revenue distribution was analyzed.

The effects of enrollment in the rural and northern school divisions should be investigated further. Partial regression indicated an effect of location on per-pupil revenue. It also indicated that students in rural areas produced declining revenues for their associated school divisions, whereas students in other locations, and especially in northern school divisions, produced increased revenues. Clearly this study was a beginning. Reasons for such obvious failures at equity enhancement should be tested. It would seem that at the least, computer modeling prior to funding revision would provide a clearer impact of funding revisions on the equity indicators used in this study. Hopefully the end result would not be just a gain in school finance theory, but more importantly, a definite gain in fiscal equity for students.

Notes

1. Telephone conversations held on October 24, 1989 in Winnipeg, Manitoba with: George Bucholz, Chairman Advisory Committee on Education Finance, Manitoba; and Denise Lovatt, Assistant Deputy Minister of Education, Manitoba.
2. Conversations with Public School Finance Board staff on October 20, 1989 and the former Chairman of the University Grants Commission in June, 1989.

References

Berne, R. & Stiefel, L. (1984). *The management of equity in school finance: Conceptual, methodological and empirical dimensions*. Baltimore, MD: Johns Hopkins University Press.

Bezeau, L. (1979). Measures of inequality of per-pupil expenditure. *Journal of Education Finance, 5*(2), 133-148.

Carruth, G.F. (1980). *Minnesota's foundation program: Projected effects of the distribution of selected maintenance revenues upon fiscal equity for fiscal years 1980-1983*. Unpublished doctoral dissertation, University of Minnesota.

Green, P.E. (1978). *Analyzing multivariate data*. Hinsdale: Dryden Press.

Hickrod, G.A.K. (1983). The decline and fall of school finance reform in Illinois. *Journal of Education Finance, 9*(4), 17-38.

Jacobson, T.P. (1986). *Minnesota's foundation program: A comparison of revenue equity between Minnesota's 1984/84 foundation program and Minnesota's 1984/85 five tier foundation proposal*. Unpublished doctoral dissertation, University of Minnesota.

Jefferson, A.L. (1982). *The fiscal equalizing effects of existing and alternative educational finance plans in Alberta*. Unpublished doctoral dissertation, University of Alberta, 1982.

Jefferson, A.L. (1985). *The Manitoba property tax credit plan: Did it achieve its intent?* (ERIC Document Reproduction Service No. ED 271 809)

Lake, P. (1983). Expenditure equity in the public schools of Atlantic Canada. *Journal of Education Finance, 8*(3), 445-462.

Lawton, S.B. (1979). Political values in educational finance in Canada and the United States. *Journal of Education Finance, 5*(4), 1-5.

MacKay, A.W. (1986). The equality provisions of the Charter and education: A structural analysis. *Canadian Journal of Education, 11*(3), 293-312.

Manitoba Government. (1959). *Final report Manitoba Royal Commission on Education*. Winnipeg: Queen's Printer.

Manitoba Government. (1964). *Report of the Manitoba Royal Commission on Local Government Education and Finance*. Winnipeg: Queen's Printer.

Manitoba Government. (1967). *White Paper on financing the single-district divisions*. Winnipeg: Department of Education.

Manitoba Government. (1981). *Manitoba education support program. A three year program of education finance*. Winnipeg: Department of Education.

Manitoba Government. (1985). *Public school financing*. Winnipeg: Department of Education.

Mendenhall, W. (1986). *Statistics for management and economics*. Boston, MA: Duxbury Press.

Schmidt, R.L. (1988). *The fiscal equalization effects of the Alberta education equity grant*. Unpublished doctoral dissertation, University of Alberta.

The Alberta Journal of Educational Research *Vol. XXXIX, No. 3, September 1993, 347-361*

Benjamin Levin
University of Manitoba

Financing Education in the Northwest Territories: Reconciling Competing Values

The literature on education finance is built around the key concepts of adequacy, equity, efficiency, and choice or liberty, which are often regarded as in competition with each other. This article uses a study of education finance in the Northwest Territories to argue that it is possible to reconcile these four values in a finance system, even where the diversity of situations and needs is great. Providing a system that pays actual costs for an agreed level of service, that leaves a great deal of local discretion as to how to use the resources provided, and that provides training and policy support to communities in making these decisions appears to be a reasonable way of combining equity, adequacy, efficiency, and choice.

Les recherches traitant du financement de l'éducation sont axées selon des concepts clés tels que la suffisance, l'équité, l'efficacité, et le choix ou la liberté, souvent perçus comme étant en compétition les uns avec les autres. En se servant d'une étude du financement de l'éducation aux Territoires du Nord-Ouest, cet article maintient qu'il est tout à fait possible de réconcilier ces quatres valeurs d'un système financier même s'il existe une grande variété de situations et de besoins différents. Pourvoir un système qui cherchera à payer les coûts actuels pour un niveau de service prédéterminé, qui permettra beaucoup de flexibilité en ce qui concerne l'utilisation des ressources disponibles, et qui offrira une formation aux communautés et appuiera les politiques de celles qui prendront ces décisions paraît être une façon raisonnable d'allier la suffisance, l'équité, l'efficacité, et le choix.

Introduction

The literature on education finance is built around the key concepts of adequacy (how much money is available for education), equity (the way the money is distributed), efficiency (the impact of money on educational outcomes), and choice or liberty (people's ability to have the kind of education they want). The meaning-in-practice of each of these has been subject to a great deal of debate, as analysts and policy makers struggle to define what these concepts might mean under particular sets of circumstances (e.g., Carnoy, 1983; Colvin, 1989). Commentators have also held that the values of adequacy, equity, efficiency and choice may be mutually inconsistent: that more of one of them necessarily means less of the others (Guthrie, Garms, & Pierce, 1988). However, recent work suggests that the relationship among the three concepts may depend on the way each of them is operationalized in any given setting. Under some circumstances, these values might be synergistic rather than inconsistent. This article uses a study of education finance in Canada's Northwest Territories (NWT) to provide a more general perspective on the ways in which a finance system can reconcile these competing values.

Benjamin Levin is an associate professor in the Faculty of Education. His main research interests are in educational policy, politics, economics, and the impact of external change on schools.

Adequacy, Equity, Efficiency, and Choice

Understandings of the meaning and implications of adequacy, equity, efficiency, and choice in education finance have been changing in important ways over the last decade. Adequacy and equity have historically been the dominating concerns. All financing systems have presumed that some basic level of schooling must be provided for there to be any benefit to students and society. Debates about adequacy hinged on what this basic level might be and how it could best be reflected in a school finance plan. Discussion of equity in education finance has traditionally been concerned with equalizing spending among jurisdictions (Barro, 1989; Paquette, 1987), a preoccupation resulting from the political division in Canada between local governments that are primarily responsible for the actual provision of schooling, provinces that are major funding sources, and national governments that have the greatest fiscal resources. The academic discussion of finance issues has paid relatively little attention to political factors and community attitudes in comparison with the realities facing policy makers.

As long as the central concern of education finance has been on resources rather than on the relationship of resources to results, efficiency was relatively unimportant (Levin, 1989; Monk, 1990; Sale, 1992). Concerns over how money was actually being used mounted as public sector resources came to be seen as limited and as evidence was advanced that increasing levels of real spending were not resulting in visibly better outcomes from schools (Easton, 1988; Economic Council of Canada, 1992; Hanushek, 1986). The debate about outcomes is a complicated one as people argue both about which outcomes ought to be valued and about how the achievement of outcomes can be assessed (Carnoy, 1983). Nonetheless, thinking about adequacy in funding now goes beyond the resources being provided in schools to encompass the impact of those resources on students.

As the belief has grown that money is not the only determinant of educational results, the relationships among adequacy, equity, efficiency, and choice concerns have become more evident. Any program of education finance is now seen to require much more than equal levels of revenue, even when some categorization of need is built into these calculations. Sensible funding requires careful attention to differences in students' needs and to the way in which available resources are actually used in schools and classrooms (Barro, 1989; Berne, 1988; Colvin, 1989). These in turn draw our attention both outward to the community context where schools exist and that shape needs and inward toward the microeconomics of resource allocation. Equity, adequacy, and efficiency are intimately associated in that situations that are equitable under traditional measures of spending can be highly inequitable in their impact on students depending on students' needs, community values, and the way resources are used. Community perceptions of equity and adequacy are likely to be more important than statistical calculations of these factors. And community perceptions of equity and adequacy are likely to be affected strongly by perceptions of efficiency.

The Northwest Territories

The Northwest Territories (NWT) comprise about 1.25 million square miles of Canada. The geography and climate are characterized by harshness. The in-

frastructure is very much based in the 60 or so communities of the NWT, most of which are small. Most transportation is by air. Some 52,000 people live in the NWT. Just under 60% of the territorial population is aboriginal—a mix of Indians, Metis (descendants of relationships between Indian and non-Indian parents), and Inuit, with the Inuit also divided into several distinct groups.

The population is young and growing quite rapidly (Statistics Canada, 1990). Some 35% of the NWT population is under 15 years of age, compared with less than 25% across Canada. The aboriginal community is increasing more rapidly than the nonaboriginal population. The NWT in 1986 had 30% of its total population in the 15-24 age bracket, compared with 21% for Canada as a whole. Population in the NWT is also growing rapidly, by more than 3% per year.

Governance

Since the mid 1960s power and responsibility have been moved to the territorial government. Power in the NWT has consequently been shifting steadily to the aboriginal population. After the most recent election in the fall of 1991, the government leader and six of the seven Cabinet members were aboriginal. In recognition of its diverse population the NWT has nine official languages, including English, French, and seven aboriginal languages.

Local government is only now in the process of development. Regional school boards have been developed in most of the NWT, but despite having elected boards the districts have few independent powers. For example, all staff, including the superintendents, continue to be territorial civil servants hired through the government of the NWT, and subject to its personnel policies. Most services to boards, such as maintenance, finance, personnel, and language translation services (especially important in the NWT where several different languages might be spoken at a school board meeting) are provided by the government of the NWT through a series of service agreements that the districts were required to sign. Schools are the property of the government of the NWT. The exception is Yellowknife where the boards operate essentially as do school districts in southern Canada.

The settling of northern aboriginal land claims and the negotiation of self-government for aboriginal groups may also have important consequences for NWT schools.

Fiscal Issues

The small population and low level of "modern" economic activity means that the territorial government has limited revenue sources. More than 80% of total revenue comes from the government of Canada through transfers. The government of Canada has been attempting for several years to reduce its own budgetary deficit, in part by reducing transfer payments to the provinces and territories, thus transferring the budget reduction decisions to another level. The NWT, with few taxation sources of its own, is particularly vulnerable to this strategy. At the same time, geography imposes high costs for many government services compared with southern jurisdictions.

Social Issues

The NWT faces serious social problems. Much, though certainly not all, of the aboriginal culture of the region has been destroyed over the past 30 years by

the encroachment of "southern" culture. Many communities suffer from serious problems of unemployment, alcoholism, and poverty. Unemployment rates for aboriginal people in the NWT in 1986 were more than 25%, compared with 10% at that time for Canada as a whole (Statistics Canada, 1990). Aboriginal people in the NWT earned in 1985 an average income of $12,000 compared with $27,000 for nonaboriginals in the NWT, and $18,000 for the Canadian population as a whole. There is a serious shortage of adequate housing in much of the NWT, exacerbated by rapid population growth (though the growth is not evenly spread across the NWT).

The overall educational level in the NWT is low, but bifurcated, with the nonaboriginal population being much more highly educated than the Canadian population as a whole, and the aboriginal population much less so. In 1986, more than 80% of NWT aboriginal people aged 15 or more, but only 28% of nonaboriginals, had less than complete secondary schooling, compared with 48% of all Canadians (Statistics Canada, 1990).

The NWT also faces significant issues of cultural conflict. Centuries-old forms of living—hunting and trapping, living off the land, and strong traditions of communal solidarity and consensus government—are colliding with the impact of television, VCRs, elections by secret ballot, the ethos of individualism, and the other accoutrements of "modern" living. One might compare the situation in the NWT in some ways to that facing many so-called "developing" countries (Poonwassie, 1991).

Education in the Northwest Territories

As of September 1989 the NWT had about 14,000 students enrolled in 76 schools. A third of these schools had fewer than five staff. The average enrollment, however, was just under 175, quite similar to rural areas of the prairie provinces. In smaller communities one school serves the entire community, extending to grade 9 or sometimes beyond. Many secondary school students must move to a larger community if they are to continue their schooling, where they live in residential facilities of various kinds. In larger communities there may be separate elementary and secondary schools with enrollments of several hundred. School facilities are often important to the community as a whole; adult education concerns are significant, especially given the generally low levels of formal education.

The NWT has low retention rates into secondary education. Only about 20% of 17- and 18-year-olds in 1988 were secondary school graduates, compared with about 65% for the country as a whole. The NWT also has by far the lowest proportion in Canada of secondary school graduates to population aged 17 or 18 (Canadian Education Statistics Council, 1990). One of the strategies for increasing retention has been to extend secondary education into more communities, but in 1990, only 8 of 60 communities provided grade 12, and only 17 provided grade 10.

Many children begin school with an aboriginal first language and little English. In some communities the aboriginal languages have themselves become corrupted by English so that students speak a language that is neither English nor the aboriginal language of their parents and grandparents.

The schools of the NWT have relatively few northern aboriginal teachers: 15% of the total in 1989, with an even smaller proportion of administrators.

Although there are several training programs in place to prepare northern aboriginal people as teachers, these are not large enough to cause the proportion of aboriginal teachers to rise rapidly. Because so many southerners are hired as teachers and do not stay in the north, turnover among teachers is approximately 30% per year across the territories. Moreover, the estimated cost of recruiting, transporting, and paying a southern teacher are about $100,000 in the first year.

At the same time, the NWT has made a firm commitment to culturally appropriate programming in its schools (Legislative Assembly of the Northwest Territories, 1982). In many areas the primary program is delivered almost entirely in the children's first language, which means that most instruction is actually done by noncertificated classroom assistants from the local community.

Education Finance in the NWT

The NWT finance system could essentially be described as an attempt to use a southern finance system, and then alter it on a case-by-case basis to meet the particular circumstances of the NWT. Given the rapid political evolution of the territories, this may have been the only strategy that was workable, although it was partly driven as well by modes of thinking that were essentially borrowed from southern jurisdictions, which is where most administrators originated.

Almost 100% of education funds are provided by the government of the NWT. The primary exception is in Yellowknife, where the two school boards levy property taxes. Yellowknife property taxes provide about 30% of local expenditure, but less than 5% of overall operating costs for education in the territories as a whole. There is almost no property tax base in the rest of the NWT, because most housing and property is owned by the government.

In traditional terms, overall per-pupil spending levels in the NWT appear higher, but not unreasonably so in comparison with southern jurisdictions given some of the diseconomies of the northern environment. Per-pupil operating spending in the NWT in 1988-1989 was $8,000, compared with about $4,500 in the rest of Canada. However, the pupil-teacher ratio in the NWT that year, 16:1, is close to the Canadian average. Moreover, expenditure per FTE student expressed as a percentage of GDP per capita is about average for the country at about 25% (Canadian Education Statistics Council, 1990). This suggests that the additional costs in the NWT are a reflection of higher personnel and operating costs, not of richer programs. The exception is in secondary education. Because of the normal organization of secondary education by subject specialty, and the small numbers of secondary students in many communities, pupil-teacher ratios in secondary education were much lower and the provision of such education therefore much more expensive.

Funding is currently provided to the school districts (not to individual schools) based on a formula plus exceptions. The formula includes a classroom staffing component based on enrollment, and additional contributions for such areas as administration, supplies, transportation, special education, operation of the regional board, building maintenance, and so on. The system has no legal mandate in legislation, regulation, or even ministerial letter. Over the years, because the districts varied so significantly, additional amounts of funding had been provided to most of them for one purpose or another. Some of

these concerned unique aspects of provision in various communities. For example, some communities required busing of students, whereas others, because of size, did not. Some communities have road access, whereas in others everything is transported by air at much higher cost. As well, almost every community has individual funding arrangements with the government of the NWT that are historical in nature, based on some particular set of circumstances that developed in a community.

Although districts are told each year how their own grants have been calculated, they are typically unaware of how other districts are funded. There is some jealousy and speculation among the districts about the special arrangements that other districts are reputed to have negotiated. And, as in most formula systems, each district feels that the formula underfunds certain key aspects of its own operations not shared by others.

Reconciling Competing Goods

The situation in the NWT illustrates the difficult nature of education finance choices. In designing a new finance system, the government of the NWT was seeking to achieve better results, to avoid large increases in expenditures, and to satisfy a diverse set of political and educational interests. The required balance had to be achieved in a setting that is significantly different from other North American situations and is in the midst of considerable social and political change. Thus issues of equity, adequacy, and efficiency were all important and interdependent.

Nor could the government of the NWT approach the problem as is typical in southern jurisdictions. Where small size or isolation are relatively infrequent, a finance system can compensate for them through categorical grants of various kinds (Monk, 1988; Verstegen, 1990). Where the bulk of students come from similar backgrounds, exceptionalities in regard to language or poverty can also be dealt with through categorical or incentive grants. Where a few schools are seen to deviate from a normal model of education, additional resources can be provided to make them more like the mainstream. Where a local tax base exists, equity concerns revolve around how state or provincial support meshes with local taxation.

None of these conditions applies in the NWT, where what is unusual in other jurisdictions is normal. Small size and isolation are general features of the system with the exception of a few schools. There is almost no local taxation capacity. Communities are quite different from each other in culture, language, and economy. And a "southern" model of education is simply inappropriate in many respects. An education finance model in the territories would have to try to deal with the issues in other ways.

A Strategy for Financing Education

How might one try to combine adequacy, equity, efficiency, and choice under these circumstances? It is only by considering both the provisions of a particular financing scheme and the link of that scheme to the broader education policy situation that one can possibly develop financing strategies that meet goals of equity, efficiency, and autonomy (Boyd & Hartman, 1988). The proposals made to the government of the NWT adopted a three-pronged approach. One element involved developing a finance formula that tried to

embody these criteria. A second important aspect involved creating the conditions under which significant local choices about educational delivery could be made. And the third element was ensuring that the finance system took into account major long-term developments in the region that would have important effects on schools.

The key elements in the school finance proposals were: providing all basic support tied to groups of students rather than classrooms; funding actual costs wherever possible; having a formula that was easily understandable; and providing funding to support local autonomy. These principles were seen as most likely to contribute to equity and efficiency.

At least as importantly, the finance scheme was tied to several broader policy considerations without which the central requirements could not be met. One central requirement was to allow the system to adjust to the devolution of authority and to help local authorities take on their increased responsibilities. A second requirement involved anticipating the major cost pressures facing the system in the next few years, so that the finance system would not become quickly outmoded.

The Proposed Financing System

The report proposed a clear financing formula, mandated through appropriate regulations, to replace the system that had grown up by accretion. Most funding, including almost all instructional funds, would be generated through calculation of the actual costs of providing an approved level of service, thus taking into account the significant variances in providing service across the NWT. Adequacy of funding would be dealt with through a base service level that would be known to all. Arguments about adequacy would then turn on whether a proposed service level was reasonable, rather than on whether a given amount of money was enough. Moreover, given the large cost differentials in different parts of the territories, a system that paid actual costs was essential to the perception of equity in funding.

Although funding was to be calculated through a formula based on the true cost of an agreed level of service, equity and choice considerations required that local authorities be able—indeed, be encouraged—to reallocate funds to provide services in the way that seemed best to them. The dilemma was to reconcile formula funding with the need for diversity in actual delivery arrangements. Where funding is provided per student or per classroom, there is pressure for actual resources to be configured in the same way as the formula, even though this may not be the most desirable or effective approach.

The solution chosen in the report was to tie most funding to groups of 50 students, a number that would force each community to make conscious decisions about group sizes and service provision. Each unit of 50 students would lead to a level of funding according to the actual salary costs of a given number of professional and support staff, with additional staff for each 50 aboriginal students. Everywhere in the NWT a group of 50 students would earn the same resources, but the dollars attached to those resources would vary depending on actual salary costs. (Because the government of the NWT bargains collectively with all teachers and support staff in the territories, boards cannot take advantage of the system by agreeing to pay higher salaries.)

All school administrators, specialists, resource teachers, aides, secretaries, and other school staff would be allocated from that single block. There would be no requirement to hire any particular configuration of staff. School and district staff and community leaders would necessarily have to think through the best allocation of staff in any given setting. Administrative support for the school districts would be provided in a similar way, as a block, based on units of 500 students. Each district would also receive an additional amount, calculated as a proportion of instructional salaries, to be used for local programming and program development. These easy-to-understand provisions would generate about 80% of the total allocation to each district.

Other provisions governed supplies costs, special needs funding, transportation costs, maintenance of buildings, administrative travel, professional development, costs of operating the divisional boards (i.e., trustee meetings and honoraria), and residences. A combination of weighted formulas and actual historical expenditures were suggested for these areas. In some areas weightings were provided to adjust base grant levels, whereas in other cases actual costs of support services were to be provided. Where there were large cost differences across the territories—such as in shipping costs, or transportation costs, or operation of residences—these differences were recognized as a matter of equity. At the same time, these acknowledgments could be made without large amounts of paperwork or the creation of an unduly complex funding formula, precisely because the bulk of the funds were generated as a block with few categories in it.

The total package would maximize the four basic values of the finance system. It would have the effect of moving the debate about adequacy toward a debate about service levels and results, rather than about percentage increases or decreases from previous years. The proposals would be equitable in that they would cover the real costs for an agreed-on level of service. They would encourage efficiency in that communities would be able to redeploy resources freely in the light of their own goals and situations. And they would accommodate, and even encourage, choice at the community level. Moreover, working toward one of the values could also foster the achievement of others. For example, the debate about what efficiency might mean and how it could best be achieved would also result in local choices about educational delivery. If the formula were seen as equitable, it might also lead to more attention to efficiency and choice issues as opposed to comparisons with others who were seen to be better off.

Support for Local Decision Making

Arguing for communities to make choices does not mean that they will be able to do so. There is some evidence that even where choice is allowed, schooling practices may not change much (Sosniak & Ethington, 1992). Community autonomy and choice involves the ability to make decisions about how best to deploy the available resources, and this in turn requires the ability to make program decisions of various kinds. Can a variety of nontraditional staff be used to deliver programs that are seen to be needed in local communities? Can adult education be given greater priority? Are local trustees able to resolve the dilemmas in regard to having language instruction provided without certifi-

cated teachers with the appropriate language skills? Must territorially determined programs be delivered in exactly the same manner in every school?

As noted earlier, the status of boards of education in the NWT was at the time of the study (and is still at the time of this writing) in flux. Boards have been created for almost all parts of the territories, but they have not been given much formal authority. The boards, and the community committees within each board, do have a certain political weight and one that could rapidly increase if boards and communities were able to mobilize citizen opinion on particular issues. At the same time, the level of political skill and sophistication is uneven. In small and remote communities, understanding of basic elements of southern Canadian politics was largely absent, and skills to deal with a large bureaucracy and a system that ran well beyond the traditional bounds of the community were not necessarily in wide supply. Thus the ability to use local autonomy to make choices would also require an investment in the skills of local people to manage (in the true sense of that word) their own education.

For local decision makers to understand and work with budgets, the finance formula cannot be unduly complex. The system as a whole should be able to be explained to intelligent laypersons and should enable them to make wise administrative decisions based on their understanding. Where reasonably competent lay people on local governing bodies cannot understand the basis of their revenue, it is unlikely that they will be able to make intelligent decisions about using that revenue. At the same time, there is a potential conflict between simplicity and equity in that a finance system that tries genuinely to respond to differences in communities and regions is likely to be made more complicated by so doing.

To support the development of local decision making, additional financial supports were proposed. Categorical funding to support the actual costs of operating regional boards and local community education committees was recommended. (It is important to remember that the costs of regional boards, in particular, could be considerable given travel requirements in the NWT. A "local" board might cover as much as several hundred thousand square miles, with communities separated by hundreds of miles of wilderness.) Recognizing that most boards would have no tax base to provide discretionary revenue, each regional board would receive an additional amount calculated as a percentage of its support under the formula, which was to be available entirely for local or regional programming priorities. As well significant targeted funding for professional development of community leaders and staff was suggested.

There were also several key areas where divisional boards and local communities had to have more authority if any education finance system was to work well. Boards would have to be able to recruit and hire their own staff, without working through Yellowknife, as do small boards all across southern Canada. Other services that had been provided entirely through the government of the NWT but could well be transferred to divisional boards included translation services, legal services, and, where boards so wished, responsibility for operation and maintenance of school buildings. Because of the costs and special difficulties of operating buildings in the Arctic, not all divisional boards at the time of the study wished to take on this responsibility.

Also important was the development of appropriate systems of planning and accountability. The Department of Education in Yellowknife would have to give up some of the detailed control it had exercised over budgets and staffing. At the same time, systems needed to be developed so that both local and territorial citizens could know how funds were being allocated and used. The best way to achieve these goals is to give specific attention to building such systems. Where there are good planning and information systems, debate can focus on substantive issues of how best to use resources, as opposed to arguing about which set of figures is correct. Developing these sorts of accountability systems is not simple. It requires thought, consultation, training, and ongoing supports. For example, the development of clear education plans by each regional board would provide an important source of accountability. But developing such plans requires skills that might need to be developed and would certainly need to be maintained and supported.

Strategic Issues

It was also critically important to take into account in the proposed finance system known factors that would, if not considered, reduce the efficacy of the system in the foreseeable future. Too many finance systems are short-term fixes to immediate problems but ignore the major cost issues facing a system in the longer term. In my view, the NWT faces major cost pressures in education from three main sources: increased expectations, high staff turnover, and increased secondary enrollments.

The main impact of increasing expectations comes from an increasing belief among aboriginal people that formal education could and should serve them more effectively than it has in past; that education can be a vehicle for their needs instead of a tool of social control and assimilation. Such expectations have implications for many aspects of school operations, and especially for curriculum, materials, and teaching strategies. Although the NWT has predominantly used the curriculum of the province of Alberta, several of the divisional boards of education were developing curricula focused on northern and aboriginal issues. Native curriculum also meant writing and producing books, especially at the primary level, written in native languages and with native content. And the demand to teach in aboriginal languages, especially at the primary level, created serious difficulties not only programmatically, but in finding staff with the needed skills.

For boards with small staff and enrollments, such as those in the NWT, these developments are expensive. For example, the Baffin Divisional Board has 3,000 students, 165 teachers, and 60 teaching assistants. The vast majority of students come to school speaking Inuktitut, the Inuit language, with little knowledge of English. To educate these students effectively, the Baffin Board has undertaken to develop an entire set of curricula focused on Inuit history and life. This in itself is a major project, especially because much of the current teaching staff does not have particular expertise in these areas. Because there were no commercial materials available in Inuktitut, the Baffin board has also created a divisional resource center that has produced (writing, illustrating, and publishing) more than 200 primary readers dealing with Inuit history, legends, and culture. Baffin has also led the move to conduct its primary program in Inuktitut, with English introduced as a second language. Few

certificated teachers in the region speak Inuktitut, so much of the primary instruction falls to teaching assistants hired from the local community. A different relationship is needed between teachers and assistants in situations where the assistants have skills that are absolutely critical to the program! In Baffin, as in the other regions of the NWT, adequacy, efficiency, equity, and choice are strongly connected.

The second major policy concern with large financial implications has to do with staff turnover. The high rate of turnover among teachers coupled with the high cost of recruiting and transporting new teachers to the north produces major expenditure pressures. At least as importantly, efficiency is reduced by having so many new teachers, so few of whom are familiar with northern communities and their needs. From both financial and educational points of view, increasing the number of northerners, and especially aboriginal northerners, in the teaching force is a highly desirable goal. And there is considerable experience in Canada, including programs in the Northwest Territories, to suggest that such a goal is both feasible and cost-effective (Unruh & Levin, 1990). However, it is not the sort of strategy that would ordinarily be thought of as an essential part of a school finance scheme, despite its important implications for finance.

A third major cost pressure comes from efforts to increase secondary school enrollments. The relatively low secondary enrollment and retention rates hid the cost of providing secondary education on a much larger scale. Given some success in reducing numbers of dropouts, and in bringing some adults back to school, it is quite conceivable that secondary school enrollments in the NWT might double or even triple in five to 10 years, with enormous cost implications. The addition of 1,500 more secondary students (doubling of 1988 enrollment) at a conservatively estimated average cost of more than $10,000 per student per year could itself account for an additional $15 million in expenditure, or a 15% increase in the total educational requirement, above any inflationary or other increases.

Moreover, there is pressure to have secondary education extended into more communities. At the time of this report, only 17 of 60 communities in the NWT provided schooling beyond grade 9. For later grades, students have to fly into one of the larger communities and live in residential facilities while attending secondary school. Given the isolation of many small northern communities and the cultural shifts involved in these moves, the system of residential schooling was widely seen as one of the causes of high dropout rates. The government of the NWT had been moving for several years to expand secondary provision in smaller communities, adding one grade at a time in several places during the 1980s. However, the cost of such provision was high, because student-teacher ratios could be as low as 3 or 4 to 1 for the secondary grades. Because the student residences in the larger communities continued to operate with high fixed costs, there were no savings from one location to pay the increased costs at the other. Clearly equity and efficiency required a different approach, and one not wedded to the conventional delivery model of one teacher on site for each subject.

Strategic direction in all of these policy areas is critical to the ability of the system to remain adequate, equitable, and effective. Some aspects of these

issues were built into the finance proposals. For example, substantial targeted funding for professional development of staff was proposed to focus on improving instructional effectiveness. Staffing for secondary schools was not enriched in the formula so as to encourage careful consideration of the tradeoffs between elementary and secondary staffing, to focus attention on alternative delivery methods for secondary education, and to encourage more sharing of resources across the levels including adult and postsecondary education. And a program of small grants to support innovation was also recommended, again as a means of fostering variety and innovation in program design and delivery. But most importantly, community choice was seen as a vehicle through which the diverse communities in the territories, with support rather than direction and control from the territorial government, would make informed decisions about how best to use available resources.

Outside the basic formula, but tied to the finance proposals, several policy directions were also proposed as an essential part of a financial strategy:

1. An investment in alternative forms of delivery for secondary education, including greater use of distance education and much more community education, is a necessity. As long as the provision of secondary education is tied to delivery in person by local teachers the costs will be prohibitive. Greater use of distance education and other technologies, combined with the development of a strong support system in the local community, offers an important and workable alternative.
2. The elimination of most residential secondary school provision was proposed, which would free funds to provide more programming in the communities, including the alternative delivery systems just mentioned.
3. A rapid increase in the proportion of northerners, and especially aboriginal northerners, in the teaching force is required through the expansion of teacher certification programs in the north, particularly for teaching assistants. The NWT already has some successful experience with training local people to be teachers, with benefits to students and communities. Investment in this area would be highly cost-effective.
4. Much stronger links are required between the public schools and the postsecondary institutions and adult education programs in the territories, making better use of scarce facilities and skilled people. The incorporation of principles of adult education into secondary and even elementary schools might well have salutary effects in its own right.
5. A strong focus should be developed on improved teaching and learning to suit the needs of the NWT. Much is known about more effective teaching and learning, but relatively little effort is devoted in most school systems to incorporating this knowledge into school practices in a lasting and productive way.

These additional, and seemingly nonfinancial, recommendations were of central importance to the overall strategy. There is nothing particularly novel about any of these proposals, but a sustained effort to address them would be a major contribution toward stability, efficiency, and effectiveness in the education finance system. Hence they formed an integral part of the overall finance strategy.

Conclusion

No system is perfect, and no claim is made that the proposals outlined would resolve all issues of education finance in the NWT. Indeed, a continuing debate over issues of adequacy, equity, and efficiency, and choice is vital if there are to be continued improvements in the allocation and use of resources for education.

At the same time, these proposals illustrate that it is possible to reconcile these four principles within a finance system, even where the diversity of situations and needs is great. Providing a system that pays actual costs for an agreed level of service that leaves a great deal of local discretion as to how to use the resources provided, and that provides training and policy support to communities in making those decisions appears to be a reasonable way of combining equity, adequacy, efficiency, and choice. Adequacy is recognized through an agreed level of services to be funded; this allows debate to focus on whether the service level is reasonable, rather than on whether the amount of money is enough (Sale, 1992). Equity is recognized through covering variations in the cost of services where such variations are essentially unavoidable. Efficiency is related in part to both the previous values. Local decision making is widely believed to enhance the efficient use of resources (Brown, 1990), while a high regard for equity allows jurisdictions to use resources in ways that make sense, instead of borrowing from one area to pay unavoidable costs in another. Moreover, using a few key building blocks for financial support allows a formula to be constructed that is both predictable and understandable.

A system that sets out basic levels of support attached to particular services but also gives significant scope for local decisions about what will actually be provided should encourage all those involved with education to debate what their schools are for and how those objectives can best be achieved. Debates about equity might focus less on whether one community has as much as another and more on what each community does with the resources available, which ties equity to efficiency. Choice is not a matter of being able to spend more or less money on the same things, but of grappling with educational provision that makes sense in a given context.

A finance system must take cognizance of the wider political, economic, and educational context. Simply to think of adequacy or equity or efficiency as issues of formula funding is too narrow; instead, it is essential to see how the wider picture will necessarily impinge on education finance and to build this strategic sense into whatever finance system is developed (Sale & Levin, 1991). In the case of the Northwest Territories, the practicality and utility of a finance system will depend on changes in the locus of political authority, improvements in management capacities, and a system that both financially and organizationally consistently promotes and supports an important local role. Finally, recognition of emerging cost pressures and provision for innovation in strategic areas through appropriate changes in service delivery are also a necessary part of meeting the challenges of education finance, in a setting as unusual as the NWT as well as in many more conventional jurisdictions.

The same issues arise in every jurisdiction. Basic systems of financing education have remained largely unchanged in Canada for many years despite

enormous changes in the surrounding social milieu—for example demographics, pressures on schools, public attitudes toward education and taxation, and so on. The strategy proposed here for the Northwest Territories could not be adopted wholesale in another jurisdiction precisely because it is designed to be sensitive to the territories' unique setting. The basic principles of attempting to reconcile adequacy, equity, efficiency, and choice can be part of a finance system in any setting. The basic approach of providing funds to meet a defined level of service while giving local option as to how to use those funds is appropriate in any setting. The need to consider education finance in the larger context of the long-term dynamics of education exists in every setting but is often relegated to a back row compared with the attention given to changes in existing formulae.

Adequacy, equity, efficiency, and choice need to be regarded as linking and complementary parts of a single approach to finance issues that is strategic and long-term in its orientation and has a strong preference toward ongoing debate about what is most desirable. It is only through an approach that integrates these various elements that we can have some realistic hope that education finance provisions might be part of improving education.

Note

This article does not necessarily represent the policy or opinion of the Department of Education of the Northwest Territories or of its staff.

References

Barro, S. (1989). Fund distribution issues in school finance: Priorities for the next round of research. *Educational Evaluation and Policy Analysis, 18*(1), 17-30.

Berne, R. (1988). Equity issues in school finance. *Journal of Education Finance, 14*(2), 159-180.

Boyd, W.L., & Hartman, W. (1988). The politics of educational productivity. In D. Monk & J. Underwood (Eds.), *Microlevel school finance* (pp. 271-308). Cambridge, MA: Ballinger.

Brown, D. (1990). *Decentralization and school-based management*. London: Falmer.

Canadian Education Statistics Council. (1990). *A statistical portrait of elementary and secondary education in Canada*. Toronto: Author.

Carnoy, M. (1983). Educational adequacy: Alternative perspectives and their implications for educational finance. *Journal of Education Finance, 8*(3), 286-299.

Colvin, R. (1989). School finance: Equity concerns in an age of reforms. *Educational Researcher 18*(1) 11-15.

Easton, S. (1988) *Education in Canada: An analysis of elementary, secondary and vocational schooling*. Vancouver: Fraser Institute.

Economic Council of Canada. (1992). *A lot to learn: Education and training in Canada*. Ottawa: Economic Council of Canada.

Guthrie, J., Garms, W., & Pierce, L. (1988). *School finance: The economics and politics of public education* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Hanushek, E. (1986). The economics of schooling: Production and efficiency in public schools. *Journal of Economic Literature, 24*, 1141-1177.

Legislative Assembly of the Northwest Territories. (1982). *Learning: Tradition and change in the Northwest Territories*. Yellowknife: Legislative Assembly.

Levin, H. (1989). Financing the education of at-risk students. *Educational Evaluation and Policy Analysis, 11*(1), 47-60.

Monk, D. (1988). Educational costs and small rural schools. *Journal of Education Finance, 16*(2), 213-225.

Monk, D (1990). *Educational finance: An economic approach*. New York: McGraw-Hill.

Paquette, J. (1987). Vertical equity in Ontario education finance. *Journal of Education Finance, 13*(2), 135-157.

Poonwassie, D. (1991). The training of substitute teachers in northern Manitoba, Canada. *Journal of Indigenous Studies, 2*(1), 43-51.

Sale, T. (1992). *The financing of elementary and secondary education in Canada*. Working Paper #29. Ottawa: Economic Council of Canada.

Sale, T., & Levin, B. (1991). Limits on the reform of education finance: A case study. *Canadian Journal of Education, 16*, 32-46.

Sosniak, L., & Ethington, C. (1992). Where public school choice is not academic. *Educational Evaluation and Policy Analysis, 14*(1), 35-52.

Statistics Canada. (1990). *Canada's north: A profile*. Ottawa: Statistics Canada.

Unruh, D., & Levin, B. (1990). Equality of access and equality of condition: Second chance programming for success. In D. Inbar (Ed.), *Second chance in education* (pp. 253-266). London: Falmer.

Verstegen, D. (1990). Efficiency and economics of scale revisited. *Journal of Education Finance, 16*(2), 159-179.

The Alberta Journal of Educational Research *Vol. XXXIX, No. 3, September 1993, 363-374*

Diana J. Nicholson

and

John O. Anderson
University of Victoria

A Time and Place for Observations: Talking with Primary Teachers About Classroom Assessment

The purpose of this study was to investigate primary teachers' use of observation as an assessment tool. Specifically, the purpose was to describe what primary teachers have been doing and what primary teachers think about what they have been doing with observation-based assessment. The study involved interviewing primary teachers to inquire about their observational assessment practices and their opinions about using observation for classroom assessment purposes. The teachers indicated that the purpose of assessment was to help students learn: assessment is a part of the larger task of education. The use of observation for classroom assessment was a predominant form of data collection with which the teachers felt comfortable. However, the use of observation was problematic. Teachers pointed out problems of representativeness of observations both in terms of the students being observed and the curricular areas involved. Participating teachers indicated that their observation efforts were not highly structured.

Le but de cette recherche était d'examiner l'utilisation de différentes techniques d'observation comme moyen d'évaluation par des enseignant(e)s du niveau primaire. Plus spécifiquement, on cherchait à décrire de quelles façons ces enseignant(e)s se servaient de ces techniques d'observation comme moyen d'évaluation et de ce qu'ils (elles) pensaient de leurs utilisations de ces moyens d'évaluation basés sur l'observation. Lors des entrevues, on invitait ces enseignant(e)s du niveau primaire à s'interroger sur leurs pratiques de l'évaluation à partir de l'observation et on sollicitait leurs opinions au sujet de l'observation comme moyen d'évaluation en salle de classe. Ils ont indiqué que le but de l'évaluation était d'aider les élèves à mieux apprendre et que l'évaluation faisait partie de la tâche plus importante qu'est la formation des élèves. Or, les enseignant(e)s ont indiqué qu'ils et qu'elles se sentaient bien comfortables avec l'observation comme moyen d'évaluation et que celle-ci servait de moyen important pour la collecte de données provenant des élèves. Cependant, leur utilisation de l'observation comme moyen d'évaluation s'avérait problématique. Les enseignant(e)s ont signalé que souvent les observations n'étaient pas nécessairement représentatives des élèves qui se faisaient observés et des domaines curriculaires traités. Les enseignant(e)s qui ont participé à cette recherche ont indiqué qu'en général, leurs observations manquaient de structuration.

The use of observation for student assessment is an activity common to many primary classrooms (Anderson & Bachor, 1992). However, little research has been conducted on this subject (Stiggins & Bridgeford, 1985). Observation has been studied most often for the purposes of evaluating its use as a research tool

---

Diana Nicholson completed her Master of Arts at the University of Victoria.
John Anderson teaches in the Faculty of Education.

(Borg & Gall, 1989; Bornstein & Lamb, 1988; Miller, 1989). Some literature cites the use of observation in educational settings, but mainly this has been for the purpose of evaluating teacher behavior (Evertson & Green, 1986; Shavelson, Webb, & Burnstein, 1986; Stiggins & Duke, 1988). The purpose of this study was to investigate primary teachers' use of observation as an assessment tool. Specifically, the purpose was to describe what primary teachers have been doing and what primary teachers think about what they have been doing with observation-based assessment.

There appear to be numerous reasons why observation has assumed a prominent role in assessment endeavors in primary classrooms. Proponents of qualitative assessment methods such as Cohen and Stern (1978) maintain that the use of observation for assessment purposes has distinct advantages that cannot be gained through the use of other assessment methods. For example, they argue that because cognitive growth is qualitative, observation is a desirable method for assessing cognitive growth because it allows for investigation of not merely the products, but also qualitative aspects of mental growth.

Similarly, Goodman (1985) states that examining children's learning through observing provides information not attainable from paper and pencil tests. Goodman contends that formal measures such as tests reflect the products of learning and the way children grow in their ability to handle testing situations without giving consideration to the importance of context of learning. Barrs (1990) declares that observational assessment enables teachers to discern patterns of learning while considering learning contexts, further allowing judgments about the ability of such contexts to meet a child's learning needs. In addition, Barrs (1990) reflects the opinion of many in the whole language movement in her assertion that standardized testing methods often strip performance down to what can be measured objectively and, by ignoring process and individuality of performance, produce information of little value to determining some important aspects of a child's development.

Bornstein and Lamb (1988) suggest that most observation occurring in the classroom probably will involve the observer/teacher determining in advance when and what to observe. They also contend that by focusing on narrow definitions of select behaviors or sequences of movements, observation can be effectively performed by observers with relatively little training in observation techniques.

The use of observation has also been promoted through program and curriculum developments. For example, the education system in British Columbia is currently involved in implementing wide-scale changes under the Year 2000 initiatives (Sullivan, 1988). The new Primary Program (BC Ministry of Education, 1990b) places considerable emphasis on the use of information obtained through teachers' observations for basing student assessment.

There is, then, support for the use of observation for student assessment purposes, but little has been done to investigate how observation is actually conducted in the classroom. As a consequence little information is available on the validity and reliability of the information collected or on the ways the information can be interpreted and evaluative decisions made. Cohen and Stern (1978) indicate that teacher education has consistently failed to include

the principles of record keeping based on observations. They suggest that teachers, therefore, frequently perceive that their records reveal little, and consequently are left to rely on hunches and intuition to form judgments. Thus it is vital for teachers to possess the ability to record and organize observations in a manner that will provide them with information that is useful for the purposes of developing learning goals and instructional plans for students.

Recommendations for developmental assessment provided by Bagnato and Neisworth (1981) reflect a variety of characteristics of observation-based assessment. For example, they suggest that assessment should consider performances across a variety of tasks in numerous settings and contexts and involve a combination of various measures and assessment information from a variety of persons in contact with the child. In addition, they contend that developmental assessment should describe actual behavior and apparent capabilities aimed at facilitating individual instructional programming. The degree to which teachers are capable or willing to follow recommended practices is largely unknown. A document on program development from BC Ministry of Education (1990a) notes that "efforts at adoption, dissemination of curriculum and materials, and professional development tell little about the extent to which teachers are using the new curriculum, how they are using it, for what reasons, and with what effects." There has been apparently little investigation into the usage of the newly outlined practices.

Although guidelines for assessment and evaluation have been provided to teachers, ultimately teachers are expected to develop and implement their own organizational systems and choose tools for use in conducting observational assessment. Consequently, decisions that teachers make will influence the degree to which their observation efforts end up being structured or unstructured, when and how they will record their observations of students, and what and how many other student assessment methods they will choose to utilize.

Conclusions from a recent study by Anderson and Bachor (1992), which described classroom assessment practices of a small sample of primary teachers in British Columbia, clearly suggest that a need persists to further examine what characterizes classroom assessment practices. Although able to state that they found assessment in primary classrooms to be continuous, individual, and integrated into classroom activities, they concluded with a list of significant questions for future research:

1. How does the teacher know what to focus on in observing students?
2. When conducting observations, how does the teacher determine which students to observe?
3. What concerns exist regarding the sampling of information and students?
4. How much information is generated and collected for each student?
5. To what extent are student behaviors comparable?
6. How representative are the performances collected of the child's repertoire?

The purpose of the study reported in this article is to identify the characteristics associated with the use of observation for student assessment purposes in primary classrooms. The research questions on which the study was based were:

1. What purposes do primary teachers feel assessment practices serve?

2. What do primary teachers believe classroom observation to mean?
3. How are primary teachers using observation techniques for assessment purposes?
4. What are primary teachers' views on observation as an assessment practice?
5. What other methods of assessment are being used in the classroom and why?

The Study

The study involved interviewing primary teachers to inquire about their observational assessment practices and their opinions about using observation for classroom assessment purposes.

A pilot study was conducted several months prior to the start of the main study in order to test procedures and the effectiveness of the interview questions developed. The pilot study was conducted in the spring of 1991; the main study began the following autumn. Pilot interviews were conducted with three primary teachers and one district helping teacher who had volunteered to be interviewed and provide feedback on the initial draft interview questions. Suggestions from these participants included revision of the wording of a few of the interview questions and the addition of five new questions. In general, consensus was that with suggested additions to the list of interview questions the interview schedule intended for use in the main study would cover adequately the major aspects of observational assessment practices in primary classrooms.

Interviewing the pilot teachers provided the researcher with a clearer vision of the problems that might be encountered in future interviews and the need to probe while questioning in order to gain more detailed responses from interviewees. Analyzing the pilot data gave the researcher some insight into the process and magnitude of the analysis that would be required in the main study. In addition, the analysis of the data obtained in the pilot study depicted four predominant themes in teachers' observational assessment practices: (a) constraints to conducting observation; (b) suggestions to improve assessment practices; (c) varied assessment philosophies; and (d) purposes of assessment. These themes assisted in refining the interview questions and in the further development of a framework for investigating the use of observation in classroom assessment.

Sample

Three school districts close to the researchers' university served as the base from which the sample of teachers was drawn. Twenty-five schools in the districts were randomly selected and the principals telephoned to request permission to approach teachers to request their participation in the study. In total, 21 teachers agreed to participate in the study. There were 19 female and two male teachers of kindergarten to grade 4, with an average of 15 years teaching experience (range: 10 to 22 years).

Procedures

Thirty-one interview questions were developed on the basis of the research questions. Interview dates and times were prearranged, either in person or by telephone, with each participant. All interviews were conducted at the school

where each participant taught. A single interview was conducted with each of the 21 participants. Of the 21 interviews, seven were conducted during a lunch break, one was conducted early in the morning when the participant had a student teacher managing the class, one was conducted during the break between a kindergarten teacher's morning and afternoon classes, and the remaining 12 interviews were conducted after school hours.

Interviews varied in length from 40 to 75 minutes. Interviews conducted during structured lunch breaks were, of course, necessarily shorter than those conducted after school hours. The schedule of interview questions developed for the purpose of guiding the interviews was generally followed during the course of each interview, although the ordering of questions posed to participants was varied according to their responses. The researcher attempted to pose interview questions in an order that related appropriately to participants' statements made in response to previous questions.

Because of time constraints, and the pattern and duration of teacher responses, not all of the 31 interview questions were consistently asked of each participant. Four of the 21 interviews did not cover all 31 interview questions. Two of the seven lunchtime interviews were shortened due to lengthy responses to initial questions that reduced the time available to cover subsequent questions. One of the lunchtime interviews was interrupted by a false fire alarm, and the interview conducted with the kindergarten teacher between morning and afternoon classes was interrupted several times by telephone calls.

Interview responses were audiotaped, and tapes were either transcribed verbatim or summarized by the researcher. Five interview tapes were randomly selected for verbatim transcription and the remaining 16 were summarized directly (Anderson, Court, & Nicholson, 1992). After producing a summary of each interview tape, summaries were mailed to participants with a thank-you note and a request for them to review and notify the researcher of any changes that should be made to better represent accuracy of meaning in their responses. Two respondents contacted the researcher with requests for minor amendments to their summarized responses. The researcher randomly selected and contacted several other participants to ensure they had received the request for review of the summary. No additional requests for amendments were made by participants.

Analyses

Analyses followed ethnographic procedures outlined in Goetz and LeCompte (1984). Although relying to some extent on the researcher's intuition and creative processes, the analyses did involve the use of systematic, formal, and logical procedures to generate categories relative to the identified research problem and to establish relationships among them.

In the initial phase of the data analysis the researcher listened to the audiotapes of the 21 interviews; in the second phase the interviews were summarized as described above.

In the third phase of data analysis, the 21 summaries of the interviews were reviewed with the intent to produce meaningful response categories for further analysis. During this review a search was conducted to identify regularities and inconsistencies in response sets. Key words and phrases

denoting general meaning in lengthy responses were highlighted for ease of referral during further analyses.

The fourth phase of analysis involved another review of individual participant summaries in which responses were categorized according to providing information to answer the five general research questions of the study. These research questions formed the typology for the analysis and the original framework for the interviews.

Three general analytic procedures outlined in Goetz and LeCompte (1984) were employed during this and subsequent phases of the data analysis. Analytic induction and constant comparison were utilized to determine which discrete responses made by individuals were relevant to each of the above stated research questions. The third procedure, enumeration, was employed during the final stage of the analysis.

Analytic induction involved scanning the data in the summaries for categories in the research questions and looking for relationships between responses and categories. During analytic induction initial cases of relevance to research questions were defined, then reviewed and occasionally modified. Using constant comparison, instances of responses relevant to the categories of research questions were compared and contrasted with other responses provided by that participant.

The fifth phase of the analysis involved the reorganization of the data produced by the previous analyses. Another summary was constructed for each participant in which their responses were organized according to the typology of the five general research questions. This was to allow for easier comparison of responses across participants in the next phase of analysis. Again, constant comparison was utilized throughout this phase of the analysis.

The final phase of the analysis involved reviewing all of the second summaries described above across all 21 participants and preparing a cumulative list of similar responses given in answer to each of the five research questions. This phase of the analysis was intended to produce a synopsis of the results that could be expressed in terms of degree of commonality of described practices and opinions. The response set for each research question was broken down into subsets of responses related to that research question, especially when responses related to the research question were many. Similar responses were enumerated across participants to represent commonly and uncommonly given statements. Negative instances of responses were included in order to qualify the extent to which responses were stated to be agreed on by participants. Responses that were excluded from enumeration dealt with issues outside the topic area of assessment and evaluation. Notations of current grade assignment and teaching experience were made under the numerated listing of participants in agreement with a specific response in order to readily identify possible trends in responses in these demographic information elements. In this final phase 13 pages of manageable data were generated from the more than 100 pages of raw data.

Results

The teachers involved in this study chose to participate and were all willing to talk about their assessment practices. They were experienced teachers. They may not be typical of primary teachers although there are no reasons to suspect

that they are not. However, in considering the results and the following discussion the reader should keep in mind that these teachers were probably operating relatively successful assessment practices in their classrooms, as they volunteered to discuss their practice with an external researcher. The study provided some answers to the five general research questions.

What purposes do primary teachers feel assessment practices serve?
Most primary teachers in the study reported believing the major purpose of assessment to be multifaceted. The most commonly stated purpose of assessment practices (by 17 of the 21 teachers) was determining how they as teachers might best help children to learn. The next most commonly stated purpose (by 8 teachers) was to provide a basis from which to communicate children's learning and abilities to parents. Other purposes noted were to provide an assessment record for future teachers, determine student learning styles, provide a benchmark for evaluation, and ensure teacher accountability.

What do primary teachers believe classroom observation to mean?
The term *observation* was stated most commonly (16 teachers) to mean simply watching students.

> Looking at how a child works and looking for relationships between child, environment, and peers.
> Looking at what children are capable of doing.

Less common themes were *listening to students* (4 teachers), recording anything children do (3 teachers), and *asking students questions* (1 teacher).

How are primary teachers using observation techniques for assessment purposes?
Interview questions pertaining to this research question generated a tremendous amount of data. Most frequently, participating teachers stated that observations were used to guide teaching and instructional goals, determine students' learning progress, and provide details and support for reporting statements.

In conducting observations, the majority of participants stated that they observed students continuously. In their responses to *when* observations were conducted, most participants distinguished between conducting informal observations *all the time* and using independent activity time, such as *center time*, or time when another teacher was present to manage the class to conduct formal observations and recordings.

In describing how they chose to focus their observation endeavors, participating teachers suggested that they generally watched the whole class and would focus on any student exhibiting behavior of interest to them. Behavior of interest was defined variously as *significant things, blatant incidents*, or things that *trigger* something. Class lists or charts were commonly utilized to ensure that all students were observed.

In addition to focusing on students, teachers would also organize their observations on learning goals for each specific grade level. Although some participants described utilizing some behavioral or developmental criteria to focus their observations, most did not declare the specifics or sources of such criteria.

The content of observation notes or records appeared to be individualized across participants. The most common notation, made by several participants, was that they included examples of *negative behaviors* in their observation notes. Although half of the respondents declared that they did a fair amount of recording mentally, thus storing observation notes in their memory, the employment of tools for recording and storing observations varied considerably across participants. Further, the times that participating teachers stated they chose to record observations differed considerably, although some form of written notes was the format most commonly adopted. Only a couple of respondents commented on the potential usefulness of a computer in the classroom. The anticipated benefits of computer use described by these respondents primarily involved the accuracy and efficiency of recording and organizing observation notes.

Although few respondents provided any sort of description of *how* they interpreted the information collected from observations, more were able to state *when* they engaged in interpreting behaviors. The majority of respondents stated that they interpreted observations both while observing and again while reviewing accumulated assessment information.

Predominantly, participating teachers described their observation practices as being relatively unstructured. Narrative descriptions of what they observed were recounted by the teachers. Reported examples of observation notes were anecdotal in nature. A common weakness of unstructured observation (Stallings & Mohlman, 1988) was frequently reflected in statements given by the teachers: the vast amount of information obtained from observations is difficult to organize and process. Participants commonly stated that record keeping requires better organizational skills than they possessed, they often found it difficult to know how to use observations for reporting, and they thought that teacher education should include training and practice in observational assessment methods. The few respondents with training in early childhood education, although among those stating a certain degree of confidence in their observational skills, agreed with the need for additional training in observational assessment.

An interesting source of observational data was observations and anecdotes provided to the teachers by other individuals. Thirteen teachers indicated that they used this information but did not elaborate on the details of the information nor its use. In general they stated that they believed it helpful to hear other people's perspectives, and they clearly implied that any information obtained from others would be weighed against their own knowledge of the child. The most common source was other teachers (9 teachers), although parents (2 teachers), and student's peers (2 teachers) were also used.

What are primary teachers' views on observation as an assessment practice?

In discussing the influence that observational assessment has on their instructional practices, participating teachers suggested benefits of use to be guiding instruction through highlighting children's abilities and learning progress, and helping to determine which activities are best suited to individual students' learning needs. In assisting them to determine learning progress, observational assessment was described as providing regular instructional feedback, paint-

ing a picture of developmental growth, and illustrating children's learning styles:

> If you give them [the children] a task and you are allowed to stand back observing, you can see that three quarters of them are missing the building blocks that will allow them to do that task.
>
> You can tell by the group responses and how the group is attacking a new activity or new learning project. If they haven't got the background then it's weak and you go back and reteach it.
>
> Observation provides more immediate feedback as to how the children are doing ... at a more regular interval.
>
> Observation provides a picture of sequential growth.
>
> You tend to individualize because of your observations ... so you have to suit your teaching to the needs of the child.

A few participants pointed out difficulties associated with translating observation information into reporting statements. Others suggested that reports using only positive descriptions about student achievement or can-do statements (which teachers inferred to be the directives from the BC Ministry of Education) may result in misleading portrayals of a child's learning progress:

> How to make observations about children that are positive and still not misleading ... we're still working on that ... that's been a big challenge and it still is a big challenge.

However, the majority of teachers viewed parents as being satisfied with the new assessment and reporting styles:

> I had a number of parents last year when I started this, come to me and say "This is the first time that I've ever read a report card where I can honestly say that you know my child."

Difficulties with predetermining the focus of observations were suggested to include periods of student activity not related to instruction or diverting attention away from students exhibiting interesting behaviors. Several respondents also agreed that it is often difficult to ensure that all students get observed; they must make a real effort to avoid neglecting to observe the quieter, less demanding students:

> That's the one failure of observation is that you have to train yourself to watch for every child. And the ones that have the problems you've got so much on , you know them inside and out. And the ones that are really bright—you know lots about those. But it's that middle group that we've always ignored in schools, and boy, in observation you find that you miss them.

In addition, several participating teachers noted that it is generally easier to get assessment information in some developmental areas than others. To this point, reasons suggested were personal beliefs about which developmental areas were more important, personal strengths in observing some developmental behaviors more readily than others, and the nature of some developmental areas being more readily observable in product rather than behaviors.

Opinions regarding recording of observations consistently reflected a lack of time for recording. Most commonly, teachers suggested time for recording required that students be engaged in independent activity and that it was difficult to develop enough activities that involve independent student activity. Some teachers suggested it was best to not worry too much about recording observations, whereas others pointed to benefits of recording such as an improved focus on students and ensuring that all students get observed.

With regard to interpreting observation information, half the respondents stated that student evaluations should be based on information obtained from numerous assessment methods. To improve the quality of information obtained from observations, it was suggested that teachers conference with students about what they have observed:

> If you observe behavior you have to ask the child about it to confirm, see what it is going on ... so I think that observation is just the beginning.

Again, time was stated to be a factor in interpretation; the accuracy of information obtained from observations required repeated observations and therefore a great deal of time.

Although it was clear that observation has been a major assessment method in primary classrooms for some time, participating teachers acknowledged some recent changes in their classroom assessment practices. Among these were: moving away from comparing students with one another, deemphasizing progress relative to curriculum while emphasizing progress relative to individualized learning goals, and moving away from "correcting" work to determining through assessment how to best help students to learn.

All but a few of the participating teachers declared that they had received no training in observational assessment during their teacher training. A few participants had received some training through their early childhood education background. Almost all the participating teachers had received some inservice training in observation. The majority stated that they perceived a need for more teacher education courses in observational assessment including more classroom experience where student teachers could concentrate on conducting, recording, organizing, and translating observations into reporting statements. In suggesting future inservice training needs in observational assessment, most participants indicated that the most productive form of inservice was to provide a forum for teachers to meet to share ideas and experiences. Additional support needs were suggested to be occasional classroom assistants to manage and guide class activities to free teachers to conduct and record observations, assistance with organizing and managing observational records, and assistance in structuring classroom activities that are more conducive to doing observation.

What other methods of assessment are being used in the classroom and why?

Participants stated that they most commonly supplemented their observations with perusal of students' work samples (19 teachers). More than half the teachers (13) declared that they found conferencing with students a useful assessment tool. Several (5) included skill checklists and self-evaluations in their repertoire of assessment tools. Three teachers reported using tests. Although several participants declared observation to be one of their preferred

assessment methods, others stated that no single assessment method was able to provide more accurate or more useful information than others; it was the combined use of a variety of assessment methods that these participants thought to be optimum.

Conclusion

The use of observation for classroom assessment was a predominant form of data collection that teachers valued and with which they felt comfortable. The teachers involved in this study indicated that the purpose of assessment was to help students learn, and that assessment was formative and intimately embedded in teaching and learning. Observation was used continuously by teachers in the classroom to provide information about student performance in order to help children learn and to report to parents.

However, the use of observation was problematic. Teachers pointed out problems of representativeness of observations both in terms of the students being observed, and the goal or curricular areas involved. Quiet students *in the middle* tended to be overlooked and some goal areas such as reading tended to receive the predominant amount observational attention. It seems that the information collected through observation would be biased toward noticeable children and particular kinds of activities. Another practice related to representativeness of data was the seemingly uncritical acceptance of anecdotes from other teachers, parents, and students as assessment related information.

Participating teachers indicated that their observation efforts were not highly structured. In general, they did not engage in assigning observations to predefined and unambiguous categories. Although most teachers said they used the developmental areas outlined in the Primary Program (BC Ministry of Education, 1990b) to guide their observation efforts, most teachers described watching students for *whatever might happen* and later attempted to match observation information to appropriate developmental areas. Although several teachers stated that they used checklists for observing simple skills, the use of checklists was not universal and was described as serving limited purposes. No teachers reported using observation schedules to plan observation times. The collection of assessment information whether by observation, portfolio, or tests should be systematically representative of relevant student abilities in important areas of learning. Procedures will have to be developed and implemented that enhance this aspect of observation as a means of collecting assessment information.

Much observation is being conducted by the teachers involved in this study, but the quality of the information and the subsequent evaluation of student achievement certainly has to be questioned. If observation is to be used in all primary classrooms to the extent that it is used by the teachers participating in this study, then there is need for the development of sound, functional, and useable procedures for classroom application.

Acknowledgment

This paper is based on the research conducted by the first author for her master's thesis. The authors are grateful for the insightful analyses and critiques of the anonymous reviewers of the first draft of this manuscript.

Note

Requests for further information on this research should be addressed to John Anderson, Faculty of Education, Box 3010, University of Victoria, Victoria, BC V8W 3N4

References

Anderson, J.O., & Bachor, D.G. (1992). What should a classroom testing program look like? The functional factors of an assessment program in primary classrooms. In D.J. Bateson (Ed.), *Classroom testing in Canada*. Vancouver: UBC Press.

Anderson, J.O., Court, D., & Nicholson, D.J. (1992). *Analysis of interview data: To transcribe or not to transcribe.* Paper presented at the annual meeting of the Canadian Society for the Study of Education, Charlottetown, PEI.

Bagnato, S., & Neisworth, J. (1981). *Linking developmental assessment and curricula*. Rockville, MD: Aspen Systems.

Barrs, M. (1990). The primary language record: Reflection of issues in evaluation. *Language Arts, 67*, 244-253.

Borg, W., & Gall, M. (1989). *Educational research: An introduction* (5th ed.). White Plains, NY: Longman.

Bornstein, M., & Lamb, M. (1988). *Developmental psychology: An advanced textbook*. Hillsdale, NJ: Erlbaum.

British Columbia Ministry of Education. (1990a). *Program development document, Learning for living—primary to graduation curriculum guide*. Victoria, BC: Queen's Printer.

British Columbia Ministry of Education. (1990b). *Primary program foundation document*. Victoria, BC: Queen's Printer.

Cohen, D., & Stern, V. (1978). Observing and recording the behavior of young children (2nd ed.). New York: Teachers College Press.

Evertson, C., & Green, J. (1986). Observation as inquiry and method. In M. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 162-213). New York: Macmillan.

Goetz, J., & LeCompte, M. (1984). *Ethnography and qualitative design in educational research*. San Diego, CA: Academic Press.

Goodman, Y. (1985). Kidwatching: Observing children in the classroom. In A. Jaggar & T. Smith-Burke (Eds.), *Observing the language learner* (pp. 9-17). Newark, NJ: International Reading Association.

Miller, P. (1989). *Theories of developmental psychology* (2nd ed.). New York: Freeman.

Shavelson, R., Webb, N., & Burnstein, L. (1986). Measurement in teaching. In M. Wittrock, (Ed.), *Handbook of research on teaching* (3rd ed., pp. 50-91). New York: Macmillan.

Stallings, J., & Mohlman, G. (1988). Classroom observation techniques. In J. Keeves (Ed.), *Educational research, methodology, and measurement: An international handbook* (pp. 469-474). Toronto, ON: Pergamon Press.

Stiggins, R., & Bridgeford, N. (1985) The ecology of classroom assessment. *Journal of Educational Measurement, 22*(4), 271-286.

Stiggins, R., & Duke, D. (1988). *The case for commitment to teacher growth: Research on teacher evaluation*. New York: SUNY Press.

Sullivan, B. (1988). *The report on the Royal Commission on Education, 1988: A legacy for learners*. Victoria, BC: Queen's Printer.

The Alberta Journal of Educational Research *Vol. XXXIX, No. 3, September 1993, 375-392*

M. Bryn Kulmatycki
Alberta Public Schools

and

T. Craig Montgomerie
University of Alberta

A Comparison of Principals' Perceptions About Their Leadership Roles in Catholic and Non-Catholic Schools

Considering that Public and Separate schools in Alberta have been funded equally since 1905, few studies have been done to compare aspects of Catholic and non-Catholic education, and even fewer to compare the roles of principals within these two dimensions. The literature indicates a perception that the role of the principal in administering schools within each of these school systems is similar. In this study a questionnaire was distributed to an equal sample of principals in both types of schools. The results of the data analysis of that questionnaire demonstrated significant differences between principals of Catholic and non-Catholic schools. Although some of the differences can be accounted for because there were no references to religious orientations on the questionnaire, some of these differences are very difficult to explain.

Même si on subventionne les écoles publiques et les écoles séparées de façons égales depuis 1905 en Alberta, peu de recherches ont été effectuées comparant divers aspects de l'éducation catholique et non-catholique dans cette province. Encore moins de recherches ont été faites comparant les rôles respectifs des directeurs et des directrices de ces mêmes écoles. Les écrits traitant de ce sujet indiquent qu'on perçoit les rôles des directeurs et des directrices de ces deux systèmes scolaires et de leur gérance comme étant semblables. Cette recherche consista d'un questionnaire distribué à deux échantillons égaux de directeurs et de directrices des deux sortes d'écoles. Or, à la fin de l'analyse des données de ce questionnaire, on a pu démontrer des différences significatives entre les directeurs et les directrices des différentes écoles, catholiques et non-catholiques. Même si on peut expliquer certaines différences en raison de la nature non-confessionnelle du questionnaire, l'explication de certaines autres différences est plus difficile.

The language used in the Alberta School Act to describe the organizational mechanics and procedures for school districts can distort a reader's orientation toward understanding the differences between Public and Separate school jurisdictions in this province. Although most districts in Alberta are either public school (Protestant) or separate school (Catholic), history has provided Alberta with a few Public Catholic school districts, and concomitantly, Separate Protestant school districts. This article maintains a distinction be-

---

Bryn Kulmatycki has been a principal in Alberta Public Schools for seven years and in Alberta Catholic Schools for seven years. His interests include concepts of administration in both Catholic and Public schools.
Craig Montgomerie is an associate professor in the Department of Educational Administration.

tween these two dimensions as simply between Catholic[1] and non-Catholic schools.

The administrative roles of principals in Catholic and non-Catholic schools in Alberta have been traditionally perceived as identical (Hochstein, 1954; Tkach, 1983; Weir, 1917). Catholic schools were not to require a distinctive administrative role or style, and the only difference between Catholic and non-Catholic schools was perceived to be the additional religious education curricular component within the Catholic school's program (Hochstein, 1955; Weir, 1917). The School Act (Alberta Education, 1988) provided both Catholic and non-Catholic school systems with identical guidelines for the establishment and operation of schools. It reaffirmed these two elements as two equal dimensions of a publicly funded education system, where it states, "Whereas there is one publicly funded system of education in Alberta whose primary mandate is to provide education programs to students through its two dimensions, the public schools and the separate schools" (Preamble).

Recent calls for educational accountability through effective schools have caused considerable examination into the principal's leadership role. Glasman (1986), and Persell and Cookson (1982) state that enhancing student achievement is a central goal of schools and the creation and maintenance of a framework for implementing this belief is critical. Rogus (1975) asserts that the principal is the most important figure in setting the upper limits for potential effectiveness of change, and Goodlad (1984) states that the principal is viewed as being the key person to effect change and improvements in the school. Burlingame (1985), Corkle (1986), and Parks (1986) state that the principal's role is varied, demanding, and requiring strong leadership, whereas Wolcott (1973) describes the principal's day as being very busy, highly unpredictable, undergoing constant change, and evolving through numerous interruptions.

This focus on the principal's role was especially timely for Catholic schools because Catholic educators began to respond to the statements on Catholic education that emerged from the Second Vatican Council (Abbott, 1966; Maycok, 1979). Since 1981 the Alberta Catholic School Trustees' Association has sponsored an annual Blueprints conference to allow Catholic school administrators to examine their role. Tkach (1983) states that these Blueprints conferences were initiated to define the role of Catholic school administrators as being characteristically different from their non-Catholic school counterparts. Laplante (1985) states that part of the Blueprints process was "to create new charters for Catholic schools that were distinct from the non-Catholic school secular creed" (p. 2). The principals of Catholic schools, then, redefined their role in terms of mission statements that advocated leadership in gospel-centered schools (Tkach, 1983).

A recent study, *The Changing Role of the Principal in Northeastern Alberta* (Department of Educational Administration, 1987), studied the status of the principalship in northeastern Alberta with respect to responsibilities, expectations, personal attributes, preparation, and qualifications and identified factors and trends influencing and reshaping the role of the principal. That study serendipitously identified that there were "dual expectations" of a principal in a Catholic school that were different from those of a principal in a non-Catholic school. Superintendents as well as board chairs indicated that additional chal-

lenges existed for principals in Catholic schools, such as additional professional development requirements for focusing on strategies to develop gospel-centered schools, and coping with changes that affect not only the educational, but also the religious component of the school program. In addition, Catholic school principals not only served as educational leaders in schools, but also functioned within a religious leadership dimension for the local Catholic community.

The primary purpose of this study was to determine if principals of Catholic schools and those of non-Catholic schools in Alberta perceived their role to be different. This study was undertaken in conjunction with another study conducted by the 1988 Administrative Development Project (ADP) of the Department of Educational Administration at the University of Alberta. That study, *The Role of the Principal: Myth or Reality?* examined aspects of the principal's role of "instructional leadership" as used in Section 15 of the Alberta School Act (1988) because "instructional leadership" is not a term defined by the Act, and no single definition is accepted or understood to be standard in educational administration literature. This study was "piggybacked" on that ADP study in order to determine if the principals of the dual-dimension education system in Alberta had similar or differing perceptions about their leadership roles.[2]

Developing a Conceptual Framework of Educational Leadership

In order to develop a conceptual framework for this study, two types of literature were examined. The first involved gathering, summarizing, and comparing the general descriptions of different leadership roles for school principals. The second involved obtaining, reviewing, and synthesizing available comparisons of these roles of principals of Catholic and non-Catholic schools.

A leadership paradigm (Figure 1)[3] was developed by graduate students of the 1989 Administrative Development Project of the University of Alberta Department of Educational Administration. A literature review revealed two groups of principals' roles that were subgrouped into a schema with four classifications. For the first group, "Instructional Leadership," tasks were classified into two leadership roles, "Direct Instructional" and "Indirect Instructional." For the second group, "Educational Leadership," tasks were classified according to the two roles, "Symbolic" and "Cultural" leadership. Some overlap occurs because some tasks are applicable to more than one classification. This paradigm was used as the basis for the data collection instrument that allowed respondents to compare attributes of all the above leadership roles against each other.

It was determined that the major difference between the role of the principal as an educational leader and the role of the principal as an instructional leader involved the extent to which the principal is directly involved with the supervision of programs, teachers, and pedagogical techniques. *Educational leadership* is defined more in terms of school unity and in developing a culture, rituals, and traditions, whereas the definition of *instructional leadership* is narrower and more measurable in terms of specific behaviors.

When compared with instructional leadership, educational leadership demonstrates behavior that is less direct in terms of physical presence. Weick

Leadership Paradigm
Leadership Role of the Principal

Instructional Leadership		**Educational Leadership**	
DIRECT MODEL	INDIRECT MODEL	SYMBOLIC MODEL	CULTURAL MODEL
Attributes Displayed by Principals for Each Type of Model			
Teacher effectiveness Student achievement Evaluation Clinical supervision Organizational vision	Established implied control Program design Climate control Organizational vision Task oriented	Unified Vision Symbolism Selective attention Purposing Universal vision Process oriented	Climate Ideology Culture Clan Bonding motivation theory Saga
Behaviors that Demonstrate Each Type of Model			
*Objectives oriented (Teacher, program, school) *Repetition of routine acts *Role-modeling good teaching *Dictates means to set goals *Concerned with increasing pupil academic achievement *Emphasizes rules and regulations *Develops and provides inservice for new curriculum *Plans program improvements *Concerned with pupil placement and class scheduling *Develops theories of instruction and learning *Models exemplary teaching *Involved in curriculum planning *Prepares master schedules	*Utilizes implicit observations *Maintains a visible presence *Provides support and resources for attainment of objectives *Emphasizes a policy approach *Diagnoses educational problems *Stresses goal setting *Provides counsel for teachers *Monitors activities in the organization *Scans events systematically *Seeks advice from staff members *Permits associates to develop master schedules *Encourages teachers to evaluate each other	*Demonstrates what is important *Communicates meaning *Creates and communicates a sense of purpose *A consensus builder *Devises a system of symbols *Builds a strong school environment *Teaches people to interpret what they are doing *Presides over ceremonies and school functions *Stresses group planning	* Emphasizes tightly structured values in a looseLY structured system. * Builds stability, certainty and predictability * Articulates school purpose and values * Uses myths that give the school a special mission or status * Articulates and celebrates emerging visions, hopes, and dreams *Creating the history of a school *Rewards accepted values and norms *Builds cultural networks *Reinvigorates rituals, traditions, and ceremonies *Defines and redefines the uniqueness of the school

Figure 1. Leadership Paradigm as developed by the graduate students in the 1989 Administrative Development Project of the University of Alberta, Department of Educational Administration.

(1982) states that a principal displays educational leadership in a more passive form through using symbols and key values. Miklos (1983) defines educational leadership as working toward implementing and improving educational programs by developing a unique organization with a culture and mission that emphasizes educational objectives.

A symbolic leader (the first subgroup of educational leadership) is a facilitator, manager, and politician who facilitates implementing and improving educational programs in the school (Miklos, 1983). Sergiovanni (1984) describes a symbolic leader as a visible consensus building leader modeling central goals and providing a clear vision for the school. A cultural leader, on the other hand, articulates a mission, develops a system of bonding symbols, and reinforces myths, traditions, and beliefs. Goodlad (1984) states that the principal provides cultural leadership through the use of developed group leadership skills, allowing each school to evolve a distinct culture and vision.

Instructional leadership is characterized by a clear commitment to academic goals that creates a climate of high expectation (Persell & Cookson, 1982). Direct leadership, the first subgroup in this category, stresses excellence and competencies (Mitchell & Cunningham, 1986); and accountability for student's performance through the principal's direct involvement in the classroom (Glasman, 1986). Indirect leadership is exercised by indirectly creating an environment where learning can thrive (Morris, Crowson, Hurwitz, & Porter-Gehrie, 1982), and by promoting the achievement of goals and the monitoring of outcomes (Manasse, 1985).

Little has been written to compare the roles of principals in Catholic and non-Catholic schools. This meant that studies of administration in Catholic schools were compared with similar studies of non-Catholic schools. This examination provided some insights into the differences between the administrators in each dimension. In addition, the religious dimension of principals of Catholic schools was examined and synthesized, and this attribute was compared with the administration of non-Catholic schools. Literature from studies done in other countries had to be examined even though Catholic schools in Alberta are not identical to the parochial schools in the United States, or to Catholic schools found in other parts of the world.

Certain obvious distinctions between these two groups of principals became apparent. Teegarden (1983) states that the role of the principal in a Catholic school is "unique," and he compares it to the role of a religious vocation. He states that Catholic school principals are additionally "called to a unique position in the Church for he or she is often in a leadership position beyond the scope of the Catholic school" (p. 68). Drahmann (1985) compared the role of the parochial school principal to that of a priest or nun, claiming that it is so important, so unique, and so pastoral that as a vocation it requires the status of a religious order.

Black (1986) and Flynn (1982) state that the attitude required by Catholic school administrators toward specific values must penetrate throughout the school, through spiritual development and establishment of an intangible force that directs the building of ethos in a school. Although the attitudes throughout the secular society often ridicule traditional lifestyles and values, and peer pressures cause the young to question the focus of Catholic schools, a

Catholic school principal must still strive to develop the students into a faith community (Eppley, 1986). McMorrow (1986) advocated a paradigm shift toward accelerated spiritual growth for lay principals of Catholic schools so that they could embrace attributes closer to those of religious orders that once dominated administration in Catholic schools.

Manno (1985) identifies pastoral and spiritual qualities that are unique to administrators of Catholic schools. Some additional roles and pressures faced by Catholic school principals, including faith commitment differences at the school site, as well as differences in the missionary vision of schools, have been identified by Pejza (1985), Fech (1985), and Dwyer and James (1985). Kelly and Metzcus (1975) found no noticeable differences between the occupational socialization and work values of parochial administrators compared to administrators of non-Catholic schools.

Few studies have been done comparing the administrative structure of Catholic and non-Catholic schools in Alberta. The legislation does not provide for any distinct administrative considerations for either of the dual dimensions (Alberta Education, 1988). Some discrepancies in the practice of administration in Catholic and non-Catholic schools are revealed when specific administrative procedures of each group are examined. For instance, in the selection process of administrative appointments, those roles with respect to the applicant's faith commitment become relevant in the selection process for lay principals of Catholic schools (Hochstein, 1955). Despite this, however, early Department of Education inspectors evaluated principals of Catholic and non-Catholic schools using the same criteria (Hochstein, 1955; Tkach, 1983; Vaughan, 1978). The Alberta Catholic School Trustees' Association (1983) identifies an additional role for the principal of a Catholic school as being a religious leader in the school community, and as being an involved religious leader in the parish community.

The study described here was designed to see if principals in Catholic schools perceived their total leadership roles as different from principals in non-Catholic schools.

The Sample

A stratified random sample[4] of 120 principals from schools operating in Alberta was selected from the Alberta Schools Database (Montgomerie, 1987). The sample was stratified by Catholic: non-Catholic; Urban: Rural; and elementary: secondary. In total 60 Catholic and 60 non-Catholic principals were selected from 21 Alberta school jurisdictions.

Method

A questionnaire containing items that compared the four elements of instructional and educational leadership qualities was developed from the Leadership Paradigm (Figure 1) and mailed to all principals. The questionnaire addressed the following five areas.

Section 1. The principals' perceptions about leadership in general were assessed by means of a series of forced-choice questions that requested that the respondent evaluate as more important their choice of each of 12 pairs of selected leadership tasks. Twenty-four selected tasks (six from each of the four leadership categories of direct, indirect, symbolic, and cultural from the

Leadership Paradigm) were paired for the respondent to select the task that was felt to be most self-descriptive.

*Section 2.*The principals' perceptions about their own leadership were obtained by asking respondents to indicate the extent that they themselves engaged in each of 24 leadership tasks on a scale of 1 (not at all descriptive of myself) to 4 (highly descriptive of myself). These items were the same tasks as used in Section 1, but were reworded to be descriptive of the respondent.

Section 3. The factors that impede principals from fulfilling a school leadership role were assessed by asking respondents to evaluate factors that block their leadership activities on a scale of 1 (does not impede at all) to 4 (impedes a great amount).

Section 4. The coping strategies for dealing with factors that impede the leadership role were obtained by asking respondents to evaluate 12 selected coping strategies that deal with factors that impeded their own leadership role. In this section a 4-degree scale of "never," "sometimes," "usually," or "often" was used to determine the level of each coping strategy that was suggested.

Section 5. The demographic information requested in this section included sex, years as a principal, total years as an educator, level of graduate education, and percentage of time formally allotted for administrative purposes.

The cover of the questionnaire contained an identification number that allowed the identification of responses from Catholic and non-Catholic schools. The number of usable responses was 88, giving an overall response rate of 77.19%. Of these responses, 45 (51%) were from principals of Catholic schools and 43 (49%) from principals of non-Catholic schools.

Results

Preferred vs. Actual Leadership Practices of Principals

Figure 2 shows the percentage of principals responding to their most preferred style of leadership.

When respondents were asked to identify their most preferred form of leadership (as opposed to the actual leadership that they practiced), no statistically significant differences emerged between principals of Catholic and non-Catholic schools.[5] In order to identify the preferred leadership style, six statements that were descriptive of each leadership style were identified. Each statement was paired with one statement from another leadership style. This led to 24 pairs of statements. Thus two statements descriptive of each leadership style were compared with each other. Respondents were asked to select a statement from each pair that they thought they would more prefer to perform. For calculation, each selection was given a value of 2 and each omission a value of 1. The maximum value for any leadership style was 12. Both groups of principals indicated that the most preferred leadership style was Symbolic, followed by cultural, direct, and indirect leadership respectively. The order of this result is inconsistent with the results comparing the respondents' description of his or her actual leadership practices, which are shown in Figure 3.

To describe their actual leadership style, each respondent was asked to evaluate six statements from each of the four categories of the leadership paradigm according to his or her own actual administrative practice. Each of the six statements was evaluated on a four-point scale from one (not at all

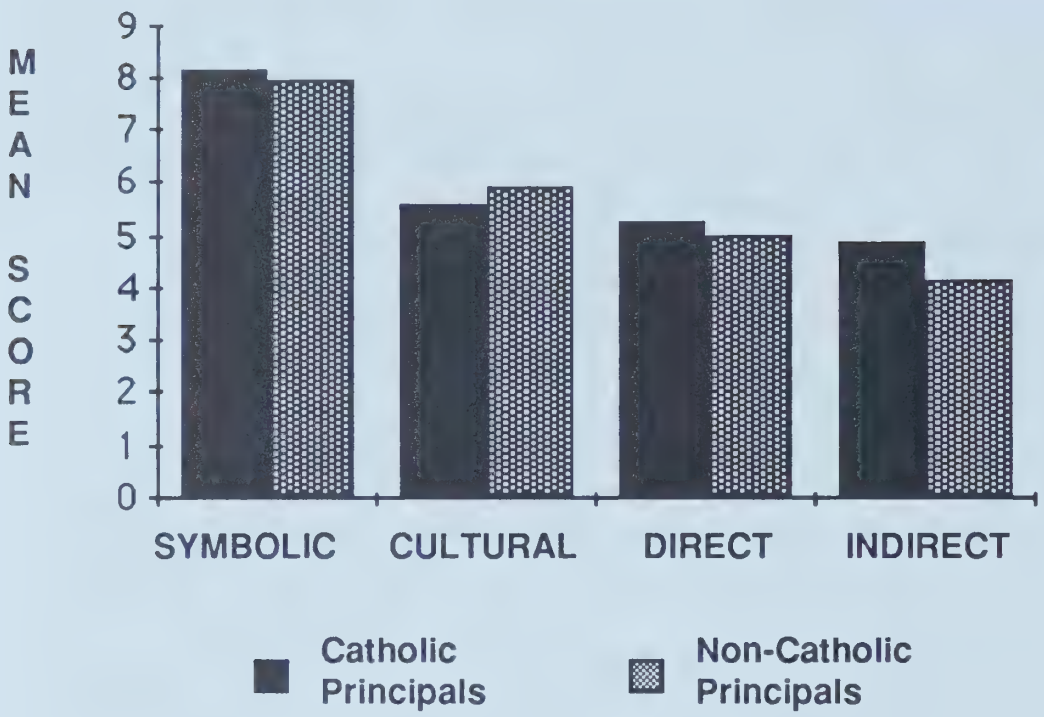

Figure 2. Preferred leadership roles as perceived by principals in Catholic and non-Catholic schools.

descriptive) to four (highly descriptive). The numbers from each of the four categories were totalled, permitting a maximum score of 24 for each category.

The responses for actual leadership indicated a different order of leadership for both groups than was indicated in the Desired Leadership of Figure 2. Further, using the 0.05 level of significance on an analysis of variance, the third category of actual leadership (indirect instructional) indicated a significant difference between the two groups ($p \leq 0.0003$). This is almost an entire descriptor level of difference between the groups. Principals of non-Catholic schools indicated that indirect instructional leadership was significantly more descriptive of themselves than did principals of Catholic schools. Both groups, therefore, find themselves doing indirect instructional leadership; however, they would both prefer to be doing direct instructional leadership, and the non-Catholic principals would prefer doing direct instructional leadership even more.

These results may have numerous explanations. Principals of Catholic schools have less administration time allocated to them, which would not allow them to engage in this time-intensive administrative style. This was shown by the response to the demographic question that indicated less administration time allocated to principals of Catholic schools (with a mean of 74.6% of their time formally allocated) compared with non-Catholic schools (with a mean of 87.8% of their time formally allocated to administration). In addition, Catholic school principals may be rejecting indirect leadership in favor of cultural or symbolic leadership, which advocates unique organizational missions, school culture, and staff motivation to buy into the mission of the school. These choices would be consistent with the ACSTA (1983) Blueprints mission statement advocating the gospel-centered mission of Catholic administration.

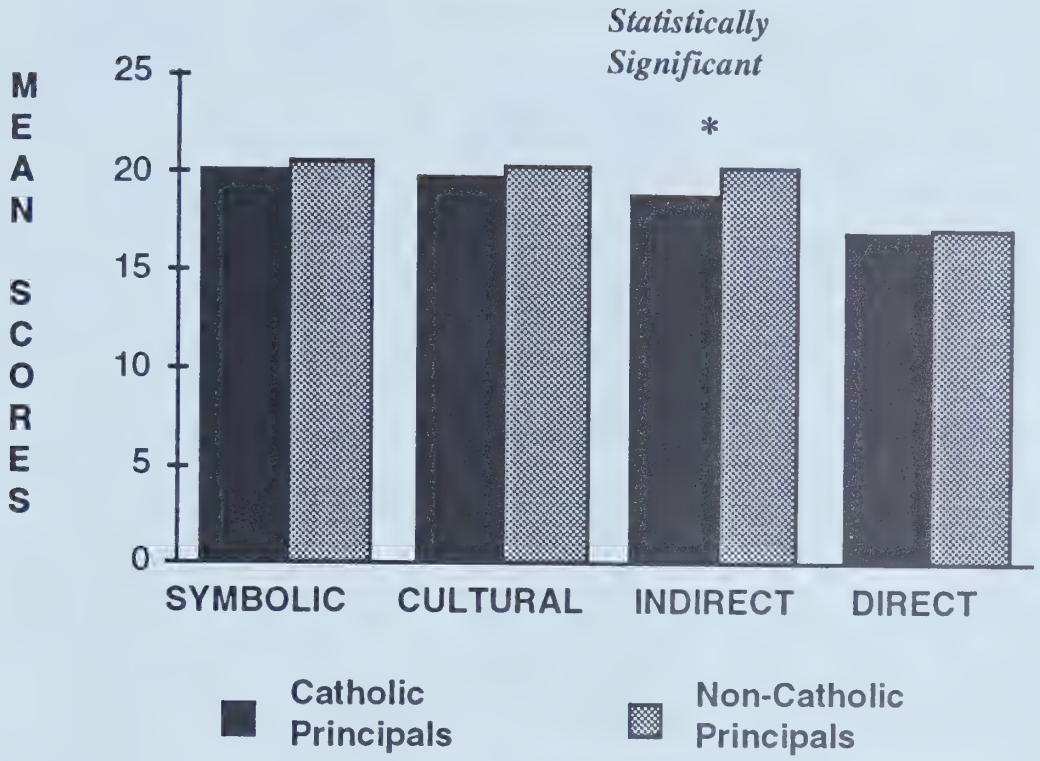

Figure 3. Actual leadership roles as indicated by principals in Catholic and non-Catholic schools.

Besides evaluating each of the four categories of leadership, the responses to individual items were also compared between principals of Catholic and non-Catholic schools, and these results are shown in Figure 4. Values were assigned to each of these 24 items such that the maximum possible score could be 4 (indicating a highly descriptive item for the respondent). The scores appeared to fall into three categories: those with an average mean higher than 3.5 (very descriptive); those with a mean between 3.0 and 3.5 (descriptive); and those with a mean of less than 3.0 (not at all descriptive).

In the first category (those with a mean higher than 3.5, or "very descriptive") one item indicated a statistically significant difference between the two groups of responses. For the item "Maintains an open door to staff for advice," the mean for principals of Catholic schools ($\overline{X}$=3.51) was significantly less than the mean for non-Catholic schools ($\overline{X}$=3.79) [t-test, p=.016]. Principals of Catholic schools perceived themselves to be significantly less open to their staff for advice than principals in non-Catholic schools. This may be a reflection of the lesser amount of administration time provided to principals of Catholic schools, which was indicated in an analysis of the demographics. The survey instrument revealed that although 40% of principals of non-Catholic schools are full-time administrators, only 28% of principals of Catholic schools were. This unavailability to staffs by principals of Catholic schools may also be a legacy from the hierarchical administrative style of the clergy administrators. Principals from religious orders may have administered schools without consultation or challenge from the subordinates within their spheres of influence, in a manner that would have been consistent with the obedience paradigm through which they were familiar with operating.

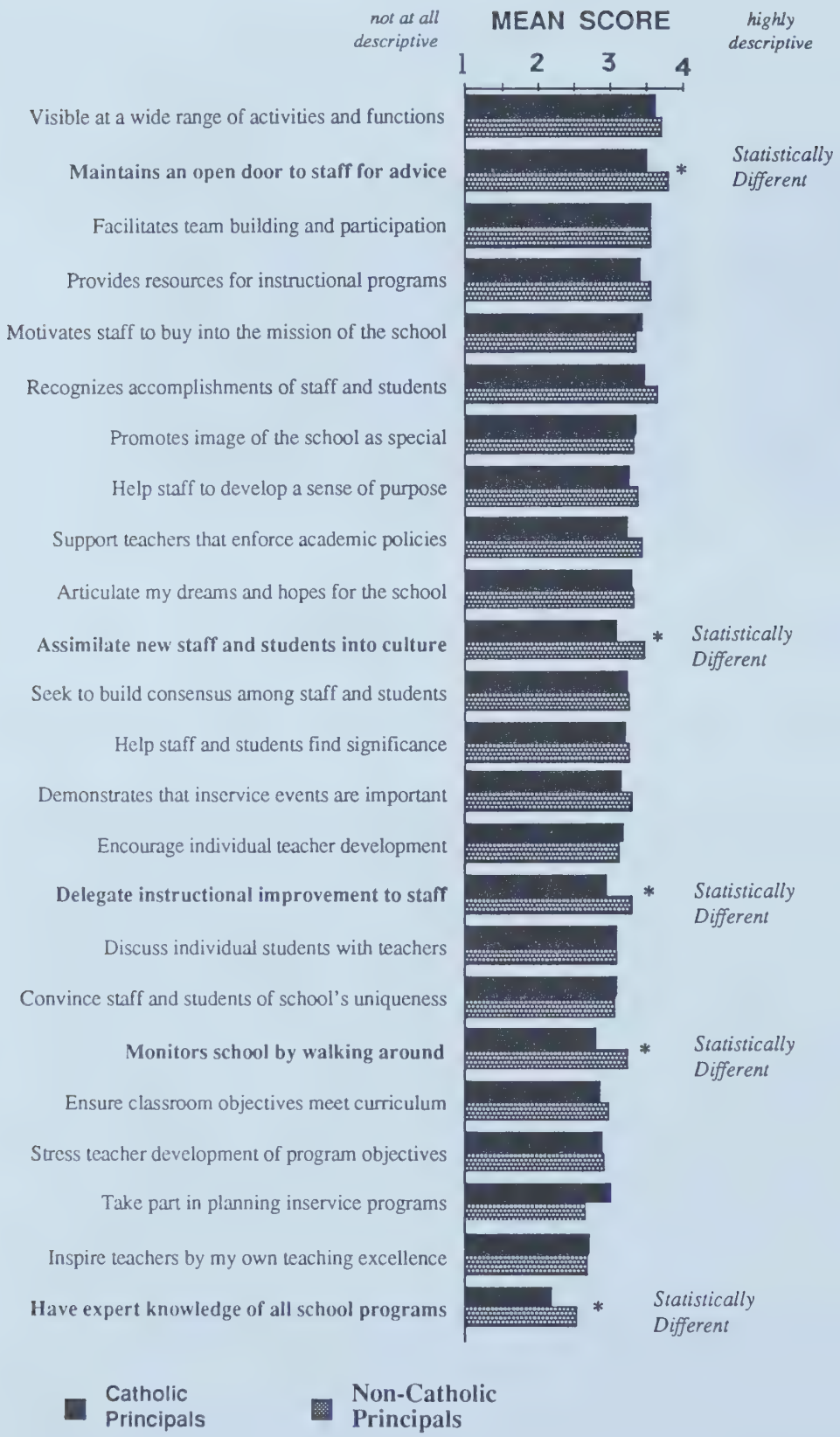
not at all descriptive
MEAN SCORE
highly descriptive
1
2
3
4
Visible at a wide range of activities and functions
Maintains an open door to staff for advice
*
Statistically Different
Facilitates team building and participation
Provides resources for instructional programs
Motivates staff to buy into the mission of the school
Recognizes accomplishments of staff and students
Promotes image of the school as special
Help staff to develop a sense of purpose
Support teachers that enforce academic policies
Articulate my dreams and hopes for the school
Assimilate new staff and students into culture
*
Statistically Different
Seek to build consensus among staff and students
Help staff and students find significance
Demonstrates that inservice events are important
Encourage individual teacher development
Delegate instructional improvement to staff
*
Statistically Different
Discuss individual students with teachers
Convince staff and students of school's uniqueness
Monitors school by walking around
*
Statistically Different
Ensure classroom objectives meet curriculum
Stress teacher development of program objectives
Take part in planning inservice programs
Inspire teachers by my own teaching excellence
Have expert knowledge of all school programs
*
Statistically Different
Catholic Principals
Non-Catholic Principals

Figure 4. Perceptions of principals in Catholic and non-Catholic schools about their own leadership roles.

The second category, "descriptive" ($3.5 \leq \bar{X} \leq 3.0$), contained three items that demonstrated statistically significant differences between the principals of Catholic schools when compared with responses of principals of non-Catholic schools. The first such item was "Assimilate new staff members into the culture of the school." The principals of Catholic schools had a mean score of 3.11, whereas the principals of non-Catholic schools had a mean score of 3.47

(t=test, p=.007). There is almost one entire descriptor level distance between the responses of the two groups. Catholic school principals, then, do not perceive themselves as assimilating new staff members into the culture of the school as much as their non-Catholic school counterparts. This result brings into question the actual nature of the Catholic school, because according the ACSTA Mission Statement Number 10, a responsibility for the principal in a Catholic school is the building of a Christian community within the school (ACSTA, 1983, p. 2). One possible explanation is that Catholic principals believe that their purpose and mission are so obvious to all new staff members that no extra effort is necessary in order to bring new staff members into the school's culture. Although this might explain the philosophical and theological aspects of the school culture, it does not explain why Catholic school principals feel they do less to incorporate new members into the culture of each individual school that has its own unique ethos. It may well be, however, that Catholic school principals do indeed perceive the culture of the school to be more important than do non-Catholic school principals. Both may actually have the same objective, but the Catholic school principals for some reason believe they should be doing more than they currently are, and therefore ranked this item lower.

The second item in this category that showed a statistically significant difference was the item "Delegates responsibility for instructional improvement to other staff members." The mean score for principals of Catholic schools was 2.96, and for principals of non-Catholic schools it was 3.3 (t-test, p=.024). Catholic school principals indicate that they perceive themselves to delegate the responsibility for instructional improvement to other staff members significantly less than non-Catholic school principals. Because Catholic schools and school districts in Alberta are generally smaller, they lack central office instructional support personnel to assist with instructional improvement. This may offer clues as to why there is such a difference between the responses. Nonetheless, it is difficult to explain why principals in Catholic schools are not delegating responsibility to other staff members, because clearly their non-Catholic counterparts are doing something very different for instructional improvement.

The third item in this category that demonstrated a statistically significant difference in response rates between the two groups of principals was the item "Monitors school operations by walking around." The principals of Catholic schools had a mean score of 2.82 whereas the principals of non-Catholic schools had a mean score of 3.25 (t-test, p=0.022). Principals of Catholic schools, then, view monitoring the school by walking around as significantly less descriptive of themselves when compared with principals of non-Catholic schools. This item may be directly attributed to the lesser amount of time allocated to principals for administration in Catholic schools as compared with the non-Catholic schools.

In the lowest category of means, which were "less descriptive of me" ($\overline{X}$<3.0), one item showed a statistically significant difference between the two groups. For the item "Displays expert knowledge of all school programs," the principals of Catholic schools had a mean of 2.18 whereas the principals of non-Catholic schools' had a mean of 2.51 (t-test, p=0.023). These results indi-

cate that although neither group of principals describes themselves as displaying expert knowledge of all school programs, the Catholic school principals indicate that they display even less expert knowledge. It is difficult to explain why there is such a difference between the responses of principals in Catholic schools and non-Catholic schools for this item.

Factors Impeding a Principal's Leadership Role

Twelve items that might impede a principal's leadership role were presented to the respondents for rating on a four-point scale ranging from level one, "Does not impede me at all," to level four, "Impedes me a great deal." Figure 5 shows the responses to these items by both groups. Catholic principals lead in all categories except two, which are "Reporting requirements of Alberta Education" and "Ambiguity in job description."

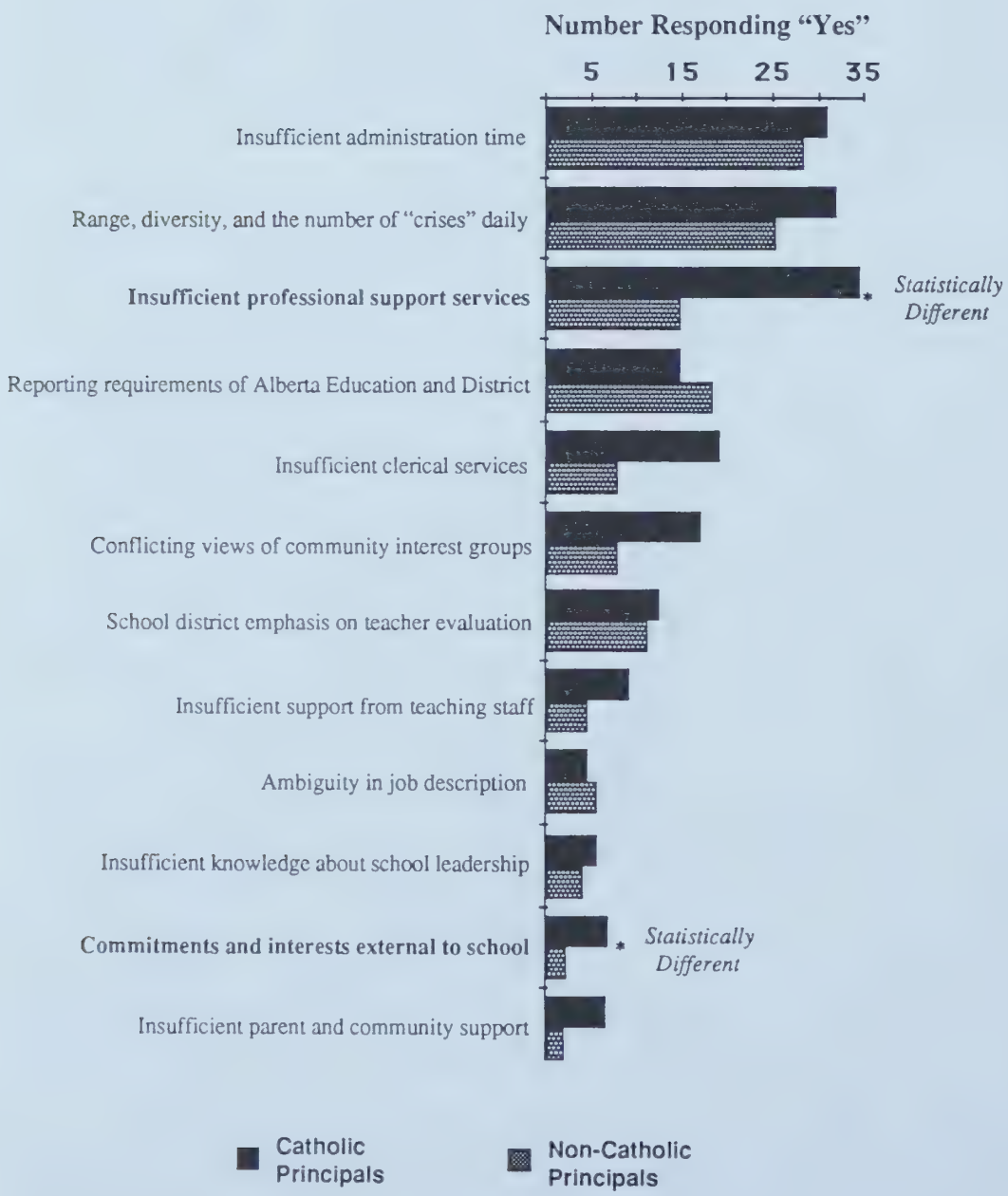

Figure 5. Factors impeding a principal's leadership role as perceived by principals in Catholic and non-Catholic schools.

The chi-square test for independence between Catholic and non-Catholic responses was performed for each item. A statistically significant dependency (p=0.0003) was found for the item "Insufficient professional support services in the school or school district (e.g., counseling)." Although both groups ranked this item the third highest of those listed as a factor that impeded leadership, the Catholic school principals (69%) perceived it to be more of a factor than non-Catholic principals (30.2%). Once again, because of the smaller Catholic districts, professional central support personnel are not available.

A second item for which a significant dependency existed (p=0.023) on the chi-square test was "Commitments and interests external to the school." Both groups rated this item low as a factor, but it was selected by 13.3% of the principals of Catholic schools compared with only 4.8% of the principals of non-Catholic schools. Initial conjecture might be that the commitments and interests to which the principals of Catholic schools are referring might be those related to the Church. This item is difficult to interpret because it is not clear from the data collection instrument as to which external factors the principals of either the Catholic or the non-Catholic schools are referring.

Coping with Factors Impeding the Leadership Role

On a four-point scale of "Never," "Sometimes," "Usually," and "Often," principals were asked to identify the degree to which 12 selected coping strategies

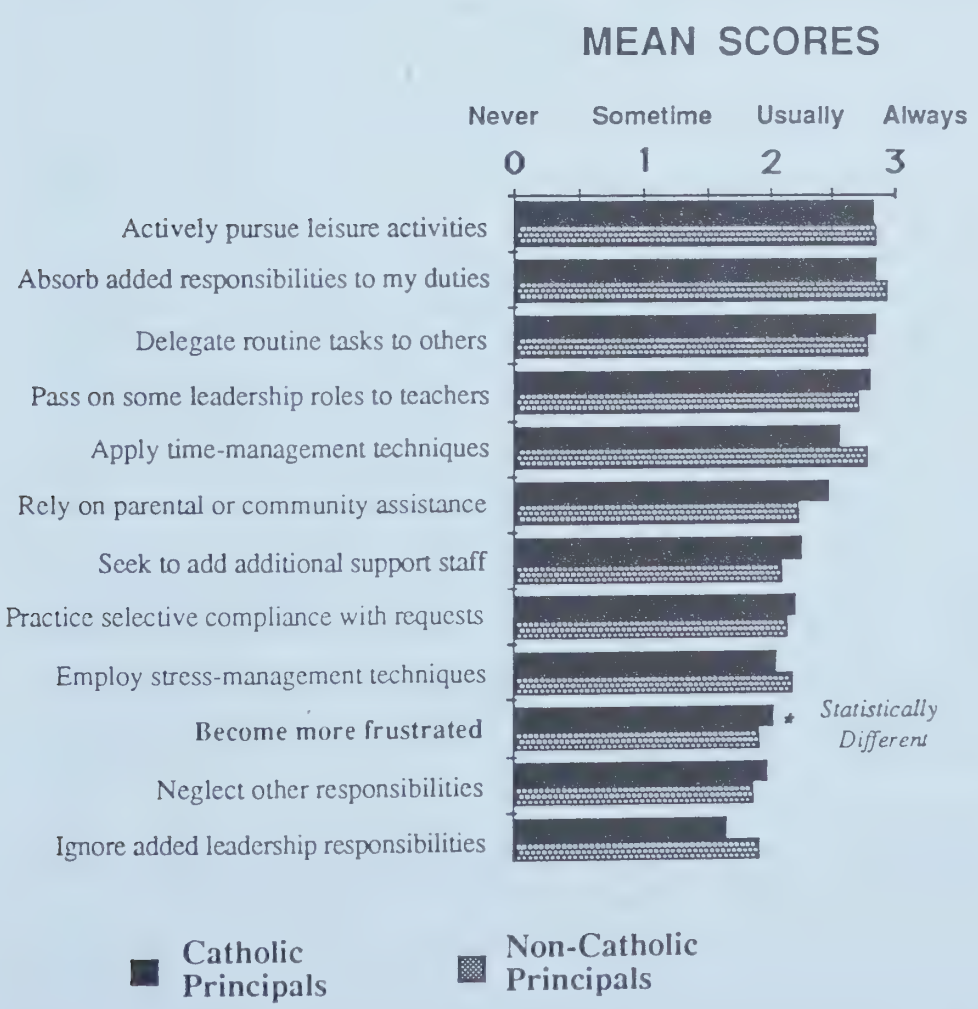

Figure 6. Mean scores for Catholic and non-Catholic principals' coping with factors impeding leadership.

impinged on their leadership role. The responses to these selected strategies are shown in Figure 6.

Principals of Catholic schools selected one item, "Become more frustrated," to a significant degree greater (t-test, p=0.045) than did principals of non-Catholic schools. The higher degree of frustration experienced by principals of Catholic schools may be the result of lesser amounts of administration time that they reported in the demographics section (28.9% of principals in Catholic schools were full-time compared with 40.5% of principals in non-Catholic schools). It may also be a result of the lesser amounts of central office support services available to principals of Catholic schools. Finally, it may be that both groups experience the same levels of frustration, except that principals in non-Catholic schools have better developed coping strategies.

Implications and Recommendations

This study has identified that although principals of Catholic and non-Catholic schools in Alberta agree in general terms about their preferred and actual leadership roles, in factors that impede the fulfilling of these leadership roles, and in coping strategies that they use, some areas of significant differences occurred between the groups.

Actual vs. Preferred Leadership Role

The preferred leadership roles were, in order: symbolic, cultural, direct instructional, and indirect instructional. When asked to describe their own leadership role, indirect instructional and direct instructional reversed order in the ranking. Second, principals of non-Catholic schools rated indirect instructional as significantly more descriptive of their own leadership role than did principals of Catholic schools. Interestingly, principals of non-Catholic schools generally responded higher on these items than did principals of Catholic schools (see Figure 3). Further, principals of non-Catholic schools related five statements significantly more descriptive of themselves than did principals of Catholic schools. These items were: "Maintain an open door to staff for advice," "Assimilate new staff and students into culture," "Delegate instructional improvement to staff," "Monitors school by walking around," and "Have expert knowledge of all school programs." Principals of Catholic schools did not rate any statements as significantly more descriptive of themselves when compared with the responses of principals of non-Catholic schools.

It is premature to suggest why these differences occur between the two groups. Future research should examine in depth whether principals of non-Catholic schools really rate themselves more positively on measures of their own leadership roles than do principals of Catholic schools, and if so, what explains this difference?

Impediments to Principals' Leadership Roles

Principals of Catholic schools were significantly higher in indicating that the lack of professional support services was a major impediment to their acting in a leadership role. Catholic school jurisdictions are generally smaller than non-Catholic jurisdictions in Alberta and not able to afford a contingent of central office personnel for curriculum and administration. This size disparity is partly because of legislation. In accordance with the Alberta School Act (1988), either the Roman Catholic or Protestant supporters (Section 199 [a][i][B]) may

establish a public school district in any area of the province (Section 194[1]). The members of the minority faith, whether Roman Catholic or Protestant, may establish a separate school district (Section 200) with the same boundaries as the public school district (Section 206[1]). In Alberta the separate school districts are generally Roman Catholic and therefore serve a minority population within a boundary determined by someone else. This predetermines that their enrollment for grant purposes would be smaller than their public school counterparts.

Smaller enrollments result in less financial support from the provincial government, making it financially difficult to employ a team of central office support services. Various administrative strategies have been attempted by Catholic school districts to respond to the need for consultative personnel with reduced cost. An option that has been used is sharing superintendency services between a number of school districts (e.g., East Central Alberta Catholic Schools). Another option has been to join Catholic administrative units together, then second their teachers to deliver consulting services to member schools (e.g., the Curriculum Resources Bank located in Sherwood Park, Alberta). Both of these options require traveling distances of hundreds of miles to assist schools.

Despite the limited availability of funds, the demands for accountability and to maintain provincial educational standards are applied to Catholic and non-Catholic schools equally, and Catholic schools must maintain credible education practices and results. In these smaller Catholic school districts, the principals may find themselves acting as the "mini-superintendent" who must handle the duties of both the principal and the superintendent in order to achieve results required by the school boards, the community, and the provincial monitoring authority.

Catholic school principals also reported that "Commitments and interests outside the school" impede their leadership role. This was rated much less important than the lack of resources, but still was significantly higher than the rating by non-Catholic principals. Previous studies (Department of Educational Administration, 1987; 1988) suggest that this may be because Catholic principals also act as a spiritual leader. It may also be because principals in small jurisdictions are required to perform tasks that might in a larger jurisdiction fall under the auspices of a central office staff member.

Finally, principals of Catholic and non-Catholic schools reported similar reliance on coping strategies for dealing with factors that impeded their ability to perform their role. Principals of Catholic schools reported "Become more frustrated" significantly more than principals of non-Catholic schools. This is probably explained by the lack of professional support services available to many of them.

Research Comparing Administration of Catholic and Non-Catholic Schools

One of the major findings of this study is that very little research has been done to compare the roles of the principals in Catholic and non-Catholic schools, or for that matter any other aspects of Catholic and non-Catholic education. As Himsl (1985) points out, "The Catholic school in Canada has not developed much that distinguishes it in administrative models and educational theory, because few persons have regarded it as a field of study" (p. 29). How, then, do

educators in Catholic schools develop research in Catholic education when almost no formal courses in Catholic school administration, theory, or history have ever existed in Alberta? This has resulted in little opportunity or encouragement for Catholic educators to explore the uniqueness of their dimension.

Conclusion

This study focuses on principals' perception of their own role and compares the responses of principals in Catholic and non-Catholic schools within Alberta's dual publicly funded school systems. It has long been accepted that the roles of the principal in both dimensions would be similar, except for the religious leadership component of the Catholic school principal's role. In fact, this study has determined that the two groups of principals do not perceive their own roles to be the same. Their differences extend into areas of educational and instructional leadership that are beyond merely the religious orientation of the Catholic school's curricular program.

It is preliminary to suggest that there are large differences between the administration of Catholic and non-Catholic schools in Alberta. This study revealed that the principals perceive their own roles differently in several areas, and it raises many issues that require further examination. Considerable study must be inaugurated in order to further verify that the Catholic and non-Catholic school systems are different enough to be viewed as administratively distinct, with needs and issues that require unique attention.

Notes

1. The definition of the term "Catholic" used in this study is based on the definition provided in Establishing Separate School Districts (Peters, 1984), in order to provide a consistent frame of reference. Peters defines Catholic in the following manner:

   > It should be noted in this context that the words *Catholic* or *Roman Catholic* are interpreted broadly to include members of the Eastern Churches, which are in communion with Rome. Altogether there are 21 different Churches in this position, though only the following 5 are officially represented in Canada: the Ukrainian Catholics, the Armenian Catholics, the Maronite Catholics, the Melchite Catholics and the Slovak Catholics. There may be individuals belonging to one of the other 16 groups however and these, if identified, should be considered, for Catholic Separate School purposes, as Catholic or Roman Catholic. (p. E-3)

2. The authors acknowledge that a substantial contribution to this study was made by graduate students in the 1989 Administrative Development Project (ADP) class of the University of Alberta Department of Educational Administration, through their project entitled The Role of the Principal, Myth or Reality? coordinated by Gordon McIntosh and Craig Montgomerie.
3. The schema entitled Leadership Paradigm (Figure 1) was prepared by the graduate students of the ADP class, and the questionnaire that was used for their study was based on this schema. Those results were reinterpreted for this study to compare the responses of principals in Catholic and non-Catholic schools.
4. In order to permit a comparison between the responses of principals in Catholic and non-Catholic schools, the random sample and data collection for the ADP study were developed and administered coincidentally with this comparative study.
5. The data for this study were delimited to the questions that were asked on the ADP questionnaire regarding the role of the principal in general. No questions regarding the religious orientation of the school, nor the religious convictions of the principal were solicited from principals on the data collection instrument.

References

Abbott, W. (Ed.). (1966). *The documents of Vatican II.* New York: American Press.

Alberta Catholic School Trustees Association. (1983). *Blueprints '83: Profiling and building the Catholic identity of our schools.* Edmonton: Author.

Alberta Education. (1988). *School Act.* Edmonton: Author.

Black, M. (1986). The Catholic school and religious education: Crucial questions for school administrators. *Catholic School Studies, 57*(2), 12-18.

Burlingame, M. (1985). *Images of leadership in effective schools literature.* Chicago: American Educational Research Association.

Corkle, M. (1986). *Improving leadership skills.* Nashville, TN: National Council of States on Inservice Education.

Department of Educational Administration. (1987). *The changing role of the principal in northeastern Alberta. A study prepared by Vermilion Outreach '87.* Edmonton: University of Alberta.

Department of Educational Administration. (1988). *The role of the principal: Myth or reality? A study prepared by the Administrative Development Program '88.* Edmonton: University of Alberta.

Drahmann, T. (1985, April). *The many-splendored role of Catholic school principals.* Paper presented at the Annual Meeting of the National Catholic Educational Association, St. Louis, Mo. (ERIC Document Reproduction Services No. ED 258 352)

Dwyer, B., & James, G. (1985). Window on a principal's life. *Catholic School Studies, 58*(1), 55-56.

Eppley, G. (1986). The role of the leadership team in Catholic schools: Australian and American perspectives. *Catholic School Studies, 59*(2), 41-43.

Fech, E.B. (1985, April). *Belief patterns of Catholic elementary school principals and teachers.* Paper presented at the Annual Meeting of the National Catholic Educational Association, St. Louis, Mo. (ERIC Document Reproduction Services No. ED 262 481)

Flynn, M. (1982). The environment and climate of Catholic schools—What now? *Catholic School Studies, 55*(1), 3-8.

Glasman, N.S. (1986). *Evaluation-based leadership: School administration in contemporary perspective.* Albany, NY: SUNY Press.

Goodlad, J.I. (1984). *A place called school.* New York: McGraw-Hill.

Hochstein, L.A. (1954). *Roman Catholic separate and public schools in Alberta.* Unpublished master's thesis, University of Alberta.

Himsl, R. (1985). The gospel-centered school. In *Visions for the Catholic School, Blueprints Reader Volume I* (pp. 29-42). Edmonton: ACSTA

Kelly, T.W., & Metzcus, R.H. (1975). Occupational socialization and work values of parochial and public school administrators: a comparative analysis. *Notre Dame Journal of Education, 6*(1), 36-42.

Laplante, R.L. (1985). *Do you mean what we mean? A Catholic school lexicon.* Edmonton: ACSTA.

Manasse, A.L. (1985). Improving conditions for principal effectiveness: Policy implications of research. *Elementary School Journal, 85*(3), 439-463.

Manno, B.V. (1985). *Those who would be Catholic school principals: Their recruitment, preparation, and evaluation.* Washington, DC: National Catholic Educational Association.

Maycok, L.A. (1979). *The influence of the Vatican Council on the American Catholic school.* Unpublished doctoral dissertation, University of Pennsylvania.

McMorrow, J. (1986). The next step: Developing the principal as spiritual leader. In *Break a leg! Blueprints reader volume II.* Edmonton: ACSTA.

Miklos, E. (1983). Alternative images of the administrator. *Canadian School Administrator, 22*(7), 1-6.

Mitchell, B.L., & Cunningham, L.L. (1986, Summer). State policy and the pursuit of instructional leadership. *Theory Into Practice,* 207-213.

Montgomerie, T.C. (1987). *Alberta schools database.* Edmonton: University of Alberta, Department of Educational Administration.

Morris, V.C., Crowson, R.L., Hurwitz, E.H., & Porter-Gehrie, C. (1982). The urban principal: Middle manager in the educational bureaucracy. *Phi Delta Kappan, 63,* 689-692.

Parks, D.J. (1986). Create a vision, build a consensus, be an effective leader. *Clearing House, 60,* 88-90.

Persell, C.H., & Cookson, P.W., Jr. (1982). The effective principal in action. In *The effective principal: A research summary* (pp. 22-29). Reston, VA: National Association of Secondary School Principals.

Peters, F. (1984). *Establishing separate school districts.* Edmonton: University of Alberta, Department of Educational Administration.

Pejza, J.P. (1985, April). *The Catholic school principal: A different kind of leader.* Paper presented at the National Catholic Educational Association Convention, St. Louis, Mo. (ERIC Document Reproduction Services)

Rogus, J.F. (1975). *Meeting the challenge of Catholic secondary education.* Calgary, AB: Religious Education Department, Calgary Catholic School Board.

Sergiovanni, T.J. (1984). Leadership and excellence in schooling. *Educational Leadership, 41,* 4-13.

Teegarden, R.A. (1983, April). Catholic school administration: A look ahead. *Today's Catholic Teacher,* 67-69.

Tkach, N. (1983). *Alberta Catholic schools A social history.* Edmonton, AB: University of Alberta.

Vaughan, K. (1978). *Catholicity and goal agreement in Catholic elementary schools having lay faculties and lay principals.* Unpublished doctoral dissertation, University of California, Los Angeles.

Weick, K. (1982). Administering education in loosely coupled schools. *Phi Delta Kappan, 63,* 673-676.

Weir, G.M. (1917). *Evolution of the separate school law in the prairie provinces.* Ottawa: Queens University

Wolcott, H.F. (1973). *The man in the principal's office: An ethnography.* New York: Holt, Rinehart & Winston.

For Contributors

Please prepare manuscripts in accordance with these guidelines.

Format

- Manuscript in triplicate.
- Typewritten or printed, double-spaced throughout.
- Abstract of approximately 100 words on a separate sheet.
- Manuscript not to exceed 6,500 words excluding graphics.
- Author's name and affiliation on a separate title page; title only on first page of manuscript to ensure anonymity in the review process.

Referencing

- Sources in parentheses after each reference (direct or otherwise), giving author's name (unless mentioned in text), year of publication, and page number(s).
- Direct quotes of 40 words or less in double quotation marks incorporated into text; those longer than 40 words in indented block format.
- All sources listed alphabetically at the end of the manuscript under the heading *References* using the APA style.
- Explanatory notes, numbered consecutively and marked in the text with superscript numerals, may appear before the References under the heading *Notes*; any citations in notes follow the same format as other references.

Graphics

- Each table and figure on a separate page at the end of the manuscript.
- Figures, charts, and diagrams in camera-ready form.
- Number tables in Arabic numerals and center over each heading, e.g., *Table 1*.
- Number figures in Arabic numerals before the caption at foot of each graph, e.g., *Figure 1: ...*
- Indicate placement of figures and tables in text, i.e., *Insert Table 3 here.*

Style

- *AJER*'s editorial style conforms closely to the *Publication Manual of the American Psychological Association.*
- Editorial changes may be made to manuscripts.
- For spelling, consult *Webster's New Collegiate Dictionary;* spelling in quoted material remains as in the original.

Diskettes

If manuscript is accepted for publication and has been prepared using a Macintosh, IBM, or IBM compatible microcomputer, please send diskette. Microsoft Word, MultiMate, WordPerfect, WordStar, Xerox Writer, XyWrite, and DisplayWrite 3 **only** are acceptable.
Important: Please indicate program used.

Submission to *AJER* implies the manuscript is not being considered for publication elsewhere.

AJER is indexed in the *Alberta Education Index, Canadian Education Index, Current Contents/Social and Behavioral Sciences, ERIC/Current Index to Journals in Education,* and *Social Science Citation Index;* appropriate articles are abstracted in *Cognitive Development Abstracts, Educational Administration Abstracts, Educational Technology Abstracts, Family Abstracts, Language Behavior Abstracts, Multicultural Education Abstracts, Psychological Abstracts, Research into Higher Education Abstracts, School Organization and Management Abstracts,* and *Sociology of Education Abstracts,* and reviewed by the *Journal for Research in Mathematics Education.*

If undelivered return to:

The Alberta Journal of Educational Research
4-116 Education North
University of Alberta
Edmonton, AB
T6G 2G5

Publications Mail Registration Number 1436